JAN 19 1959

CHILD DEVELOPMENT

Vol. 29
December, 1958
No. 4

Published Quarterly by

CHILD DEVELOPMENT PUBLICATIONS

OF THE

SOCIETY FOR RESEARCH IN CHILD DEVELOPMENT, INC.

Contents for December

Adult Aggression in the Management of Disturbed Children
Elton B. McNeil and *J. Robert Cohler, Jr.* 451

A Self-Concept Scale for Children and Its Relationship to the Children's Form of the Manifest Anxiety Scale
Lewis P. Lipsitt 463

Size Judgment at Different Distances as a Function of Age Level
Walter Cohen, Aaron Hershkowitz, and *Marjorie Chodack* 473

Developmental, Psychosocial, and Educational Factors in Children with Nonorganic Articulation Problems
Ruth FitzSimons 481

Self-Conceptions, Motivations, and Interpersonal Attitudes of Early- and Late-Maturing Girls
Mary Cover Jones and *Paul Henry Mussen* 491

A Reinterpretation of the Barker, Dembo, and Lewin Study of Frustration and Regression
John M. Davis 503

Dominance Attitudes in Parents and Adult Avoidance Behavior in Young Children
E. Philip Trapp and *Donald H. Kausler* 507

A Study of the Validity of Parent Attitude Measurement
Donald A. Leton 515

Developmental Status of Two Groups of Infants Released for Adoption
D. Bruce Gardner and *Marybelle K. Swiger* 521

Variables in Early Discrimination Learning: II. Mode of Response and Stimulus Difference in the Discrimination of Tonal Frequencies
Wendell E. Jeffrey 531

CHILD DEVELOPMENT

Vol. 29
December, 1958
No. 4

Editor

WILLIAM E. MARTIN
Purdue University

Consulting Editors

BETTYE McDONALD CALDWELL
Washington University School of Medicine

ERCEL S. EPPRIGHT
Iowa State College

EUGENE S. GOLLIN
Queens College

HARRY LEVIN
Cornell University

PAUL H. MUSSEN
University of California

KATHERINE E. ROBERTS
Merrill-Palmer School

WATARU W. SUTOW
University of Texas

WILLIAM W. WATTENBERG
Wayne State University

Managing Editor

KATE HOFFMAN

Published Quarterly by

CHILD DEVELOPMENT PUBLICATIONS
(of The Society for Research in Child Development, Inc.)
Purdue University
Lafayette, Indiana

Made in United States of America

Notice to Contributors

Child Development is devoted to basic research and theory concerning the growth and development of the child.

Manuscripts should be submitted in duplicate, each copy bearing only the title as a means of identification. The title, author's name, and institutional affiliation (together with any footnotes which reveal the identification of the author) should be submitted on separate pages.

Contributors are asked to refer to the *Publication Manual* of the *American Psychological Association* (1957 Revision) in the preparation of manuscripts.

The following rules require special emphasis:

1. All copy, including indented matter, should be typed double-spaced on white standard paper. Lines should not exceed six inches in length.

2. Footnotes should be few in number. They should be indicated in the text by superscript arabic numerals, numbered consecutively throughout the manuscript. The footnotes themselves should be typed double-spaced on a separate page placed at the end of the manuscript.

3. Tables and figures are costly to print and should be used only when essential. The author will be asked to contribute to the cost of composition of tabular and graphic material, which is excessive in amount or unduly complicated in form. Each table should be typed on a separate sheet. Insert a guide line (e.g., Table 1 about here) at the appropriate place in the manuscript. Figures should be prepared with black India ink, suitable for photography. Legends typewritten on an ordinary typewriter are unacceptable.

4. References in the text should be indicated by numbers in parentheses that refer to the list of references at the end of the article. Do not use footnotes for this purpose. In the list of references, the arrangement should be alphabetical by author's surname, or for multiple authors by the name of the senior; for any one author, by order of date. The following standard form should be observed:

> Jones, H. E. Development in adolescence. New York: D. Appleton-Century, 1943.
>
> Shepard, Winifred O., & Schaeffer, M. The effect of concept knowledge on discrimination learning. Child Develpm., 1956, 27, 173-178.

The World List of abbreviations is used for journals.

The manuscript submitted should represent the final form in which the author wishes the paper to appear. All changes in proof (other than printer's errors) will be charged to the author.

All manuscripts should be addressed to William E. Martin, Editor, Child Development Publications, Purdue University, Lafayette, Indiana.

ADULT AGGRESSION IN THE MANAGEMENT OF DISTURBED CHILDREN[1]

ELTON B. MCNEIL
University of Michigan

and J. ROBERT COHLER, JR.
Harvard University

In a situation where care must be provided for aggressive children it is imperative to select workers whose control of their own aggressive impulses is adequate to the task. Essentially, a judgment of the worker's potential behavior *in an aggressive situation* is called for and most institutions must either make this selection on the basis of unreliable hunches or rely on a trial period on the job to test the mettle of employees or trainees. The demonstration of a predictable relationship between test measures of personal hostility and ability to manage hostile charges effectively would be of considerable value in preselection of persons for this kind of demanding work.

There have been a number of explorations of the elusive relationship between fantasy and actual behavior. Studies (1, 2, 3, 4, 5, 6, 10, 11, 12, 13) of the connection between hostile fantasies and aggressive actions have dealt with populations ranging from delinquent boys to psychotic adults and have employed a number of measures, such as the TAT, doll play, and story completions, to assess the psychic activity of the individual. Without discussing the details of the findings in each study, it seems clear that the assessment of actual aggressive behavior on the part of the subject has been a chronic problem. Doll play, for example, can simulate a real life setting, but it remains, at best, a miniature and highly symbolic instance of free emotional expression. The fact that the doll victims of aggressive outbursts are inanimate and incapable of responding in kind places a severe limitation on the generalizability of these findings to other more complex

[1] This research was supported by a grant from the Social Science Research Council.

settings. The next most frequent method of determining how much aggression subjects normally express has been by means of tests of reputation. Ratings and rankings of past behavior are gathered from knowledgeable observers and used as criterion measures. In many instances, the reliability of the raters is an unknown quantity, and there has been no opportunity for raters to observe all of the subjects responding to a similar set of emotional pressures. Without some standardization of the situation in which ratings of aggression are made, the meaning of high or low aggression becomes equivocal. Methodological difficulties of this sort have prevented researchers from making a definitive statement regarding the relationship of aggressive fantasy to actual behavior.

For our purposes, the most important theoretical limitation of such studies is that their findings apply particularly to those everyday human relations in which open and violent aggression seldom occurs. Often, it is important to be able to predict human behavior when faced with the stress of undiluted resistance, defiance, or attack. The relationship of fantasy to behavior, under such circumstances, has yet to be fully explored. While the degree of hostility present in a person's fantasy life may not be predictive of loss of self-control in the course of normal human affairs, it might well be an accurate index to his response under substantial interpersonal duress.

The present study is an attempt to explore the relationship between two test measures of hostility and the ability to manage personal aggression in a therapeutic camp for disturbed children. A general hypothesis to be tested is that success in the management of aggressive children bears a predictable relationship to the degree of hostility present in the fantasy productions of the adult child care worker.

Method

Research Setting

The University of Michigan Fresh Air Camp provides an ideal setting for field observations of hostility and conflict. Under the guidance of a trained clinical staff and 40 student counselors, the Fresh Air Camp conducts "two, month-long terms of mental-hygiene oriented camping for 120 seriously maladjusted boys, ranging in age from 7 to 14 and referred for this summer experience by one of some 55 treatment agencies" (8). The majority of these children act out their conflicts and openly express hostility toward peers and adults. Throughout the day the counselor is confronted with a series of aggressive incidents requiring intervention and therapeutic handling. The counselor's ability to control his own hostility, as well as the hostility of others, is being tested constantly.

The subjects in this study were 18 male and 15 female counselors who participated in the 1955 session of the Fresh Air Camp. The mean age of the male counselors was 28 while that of the female counselors was 23.

Measures of Hostility

During an orientation and training period at the beginning of the camp session, two measures were used to provide a means of ordering the counselors along a dimension of personal hostility and aggression. The first measure, a test constructed by Watson, Pritzker, and Madison (13), consists of 60 sets of scrambled words which can be assembled to have either a hostile or neutral meaning. For example, the set of words, "Shoot I'll you ask," may be assembled as "I'll ask you" or "I'll shoot you" by omitting either the neutral or the hostile word. The 60 sets of words represent a wide range, in degree of intensity, from very mild to very strong hostility. Watson, Pritzker, and Madison found that these scrambled sentences could be used successfully to differentiate between a neurotic patient group and a group of control subjects with the neurotics showing significantly more hostile constructions than the normals.

A second measure of hostility was derived from the scoring of spontaneous stories elicited in response to the presentation of four TAT (9) cards (3 BM, 8 BM, 18 BM, and 18 GF). A study by Holzberg, Bursten, and Santiccioli (4) offers experimental support for the proposition that individuals whose stories contain an unusual response (greater or less than average) to the aggressive implications of situations suggested by these TAT cards are more highly charged with aggressive tension than are those who respond with an average number of aggressive stories. In their study, aggressive tension was measured in terms of the relative efficiency of learning aggressive, as compared to neutral, material.

Counselor Sociometric Ratings

Working together in the psychologically intimate setting of a therapeutic camp, the counselors soon make judgments regarding one another's personal characteristics and therapeutic skill. These peer judgments often are merciless in their accuracy, for they reflect two months of daily observation of one another's actions in a variety of circumstances. At the conclusion of the camp session each counselor rates the other counselors along a series of dimensions ranging from choice as a personal friend to potential as an understanding parent. For purposes of this study three questionnaire items were chosen as most relevant in estimating counselor reputation for managing aggressive and disturbed children.[2] For each item the rater was asked to indicate the three counselors who most displayed the characteristics as well as the three who least displayed them. After summing the frequency with which each counselor was nominated as possessing the

[2] ITEM 3: If I were director of the Fresh Air Camp and given the opportunity of increasing the staff, I'd prefer (*3 most*) (*3 least*). ITEM 6: If I were to choose the outstanding counselors, from the standpoint of being understanding and effective with the children, I'd prefer (*3 most*) (*3 least*). ITEM 9: The counselor who is best controlled with his children, doesn't lose his temper easily or employ coercive and aggressive methods in dealing with discipline problems is (*3 most*) (*3 least*).

most or least of each characteristic for all three items, the authors converted the scores to percentages and rank-ordered all counselors. The rank-order distribution was divided into three equal groups of persons designated as Most Chosen, Average Chosen, and Least Chosen on this measure of counselor preference.

Counselor Behavior

In dealing with disturbed children, everyday situations such as eating, sleeping, or playing are always potential settings for aggressive explosions. Under the guise of interest in determining the causes of such outbursts among the children, we enlisted the aid of 15 counselors who permitted us to observe their interactions with their cabin group. During the camp session observers recorded 10 aggressive incidents from the beginning to end for each of 15 volunteer counselors. The descriptive accounts of these incidents were coded along two dimensions: (a) the effectiveness with which each situation was resolved; (b) the degree of acceptance, by the child, of the solution.

Results

Measures of Hostility and Sociometric Choice

In order to test the relationship between our measures of aggression and the counselors' sociometric ratings of one another's ability to manage aggressive child behavior, chi square analyses were made using the Mood Likelihood Ratio Test (7). The range of scores on the Scrambled Sentence Arrangement Test (from 5 to 27 hostile completions) was divided into three groups—High Aggression, Medium Aggression, and Low Aggression —according to the frequency with which the words were assembled to have a hostile meaning. For the four TAT stories the scores varied between no aggressive stories to four. This restricted distribution necessitated the division of responses into two groups, High Aggression and Low Aggression.

The scores on the Counselor Sociometric Choice measure of ability to handle hostility in disturbed children were characterized as Most Chosen, Average Chosen, and Least Chosen. Our Most Chosen counselors were those judged by their fellow workers to possess the greatest facility in managing aggressive youngsters while the Least Chosen were those judged to have had less ability to manage such children successfully. Comparisons of counselor scores on the test measures of hostility and the peer rankings of skill in managing the children are contained in Table 1. The TAT measure of hostility bears a significant relationship to sociometric choice with the Low Aggression subjects the Most Chosen as successful by their peers while the High Aggression scores are to be found among Average or Least Chosen counselors. Although the direction of the cell loadings is similar to the TAT, the Scrambled Sentence Arrangement Test of hos-

TABLE 1

THE RELATIONSHIP OF COUNSELOR SOCIOMETRIC CHOICE TO COUNSELOR AGGRESSION ON THE TAT AND THE SCRAMBLED SENTENCE ARRANGEMENT TEST *

	Sociometric Choice				
	Most Chosen	*Average Chosen*	*Least Chosen*	χ^2	*p*
TAT					
High Aggression	2	8	5	6.99	$< .05$
Low Aggression	9	3	6		
Scrambled Sentences					
High Aggression	3	6	2		
Medium Aggression	3	4	4	6.07	$< .20$
Low Aggression	5	1	5		
Combined TAT and Scrambled Sentences					
High Combined Aggression	1	6	1		
Medium Combined Aggression	4	5	5	14.96	$< .01$
Low Combined Aggression	6	0	5		

* In this table, as in the others to follow, two-tailed tests of significance were used. Although prediction of direction of cell loading was made for each table, the small *N* dictated the use of the more rigorous statistical model required by a two-tailed rather than a one-tailed test of level of significance.

tility does not display a significant relationship to the Counselor Sociometric Choice measure.

In an attempt to devise a better measure of hostility, the scores on the TAT and the Scrambled Sentence Arrangement Test were combined to produce a single measure. This was accomplished by making a new rank-order distribution of the raw scores on the TAT and Scrambled Sentence Arrangement Test. Merging these two distributions into a single measure of Combined Aggression and dividing this measure into three equal groups of persons, categorized as High, Medium, and Low, required some decision making. At the high and low extremes, for each measure, categorization posed no problem. In those cases in which a counselor scored high on one measure and low on another he was relegated to the Average category, since designation as either high or low on the Combined Aggression measure could not be defended. This procedure allows the extremes to become fairly pure cases while the Average category is composed of persons average on both measures or high on only one. Categories of High, Medium, and Low Combined Aggression were formed and compared with sociometric choice. This combined measure reveals that a High Combined Aggression

score is most regularly associated with the Average Chosen category, while Medium Combined Aggression scores appear equally in the Most, Average, and Least Chosen categories. A Low Aggression rating is distinctly related to being either Most or Least Chosen by others.

Hostility, Sociometric Choice, and Child Management

A clearer understanding of the meaning of the counselors' sociometric choices of one another can be derived from an examination of their relationship to actual counselor behavior with the children. The expert observations of counselor handling of conflicts between children were coded according to the effectiveness with which these conflicts were resolved. Effectiveness, in this instance, was judged in terms of the degree to which peace was restored in the group, the time required to negotiate the cessation of hostilities, the permanence of the resolution, and the need for help in achieving a solution. On the basis of a cumulative score from 10 such incidents, each counselor was classified as High, Medium, or Low in Effectiveness.

In many instances an uneasy peace was established at the expense of the child who had initiated the conflict. Although such instances might reasonably have been included as an indication of ineffective conflict resolution, it was a sufficiently frequent and distinct characteristic of counselor-child relationships to warrant special consideration. The child's perception of his treatment and his behavior following the counselor's attempt at resolution were coded into categories of High, Medium, and Low Acceptance.

Table 2 reveals that both conflict resolution and the child's acceptance of the resolution bear similar and highly significant relationships to sociometric choice. Our Most Chosen counselors regularly are found to be very effective in resolving conflict and have a high acceptance of these resolu-

TABLE 2

THE RELATIONSHIP OF COUNSELOR SOCIOMETRIC CHOICE TO CONFLICT RESOLUTION AND CHILD ACCEPTANCE OF RESOLUTION

	Sociometric Choice				
	Most Chosen	*Average Chosen*	*Least Chosen*	χ^2	*p*
Conflict Resolution					
High Effectiveness	4	0	1		
Medium Effectiveness	0	4	1	13.44	$< .01$
Low Effectiveness	1	1	3		
Child Acceptance of Resolution					
High Acceptance	4	0	1		
Medium Acceptance	0	3	2	10.67	$< .05$
Low Acceptance	1	2	2		

tions by the child. The reverse of these findings is suggested for our Least Chosen counselors. Thus it seems that the choice of successful and unsuccessful counselors by their peers is an accurate reflection of their handling of difficult situations, as judged by an outside observer.

Sociometric choice, then, relates significantly to the TAT measure of aggression, the combined measure of aggression, the counselor's ability effectively to resolve conflict, and the child's acceptance of the resolution of the conflict. An important next step is to determine the nature of the relationship between our precamp measures of counselor aggression and the later behavior of the counselors in dealing with hostile children.

TABLE 3

THE RELATIONSHIP OF COUNSELOR AGGRESSION ON THE TAT TO CONFLICT RESOLUTION AND CHILD ACCEPTANCE OF RESOLUTION

	TAT			
	High Aggression	*Low Aggression*	χ^2	*p*
Conflict Resolution				
High Effectiveness	1	4		
Medium Effectiveness	5	0	8.46	$< .02$
Low Effectiveness	3	2		
Child Acceptance of Resolution				
High Acceptance	1	4		
Medium Acceptance	4	1	5.18	$< .10$
Low Acceptance	4	1		

Table 3 indicates that a significant relationship exists between the TAT measure of aggression and Effectiveness in Conflict Resolution, but that the association with degree of Acceptance of Conflict Resolution by the child attains only a borderline level of significance. Counselors who were rated Low in aggression on the TAT tended to be more effective in resolving hostility and in having the child feel that he was dealt with reasonably and fairly. In general, counselors rated High in aggression received no more than a Medium or Low rating of effectiveness and found a similar degree of acceptance by the children. In all fairness it ought to be noted that the category designated as High Acceptance of Resolution was quite relative to the group and the setting. Although the degrees of acceptance were distinguishable from one another, their over-all range was somewhat restricted. It was seldom possible to achieve enthusiastic acceptance from our revenge-motivated children, and this fact made the task of prediction a particularly difficult one.

The scrambled sentences measure of aggression, which failed to bear a significant relationship to counselor sociometric choice, was similarly without predictive value when compared with either Conflict Resolution or Child Acceptance of Resolution.

The combined test measure of counselor aggression also failed to differentiate counselors in terms of the effectiveness of their attempts to resolve conflict and in the acceptance of the resolutions by the children.

Discussion

The theoretical and methodological problems of assessing the relationship of aggressive fantasy to hostile behavior are apparent in the results of this study. Table 1 underlines a number of these problems because it contains implications for both the assets and limitations of such efforts. Both the TAT and the Combined TAT and Scrambled Sentences measures are able, empirically, to predict successfully to peer ratings of clinical skill; yet this finding leaves us adrift theoretically. It is clear that the prediction of behavior from fantasy with these tools is useful as a wide screening device but does not provide answers to our fundamental questions regarding the meaning of hostility in the individual's psychic economy. In particular, we have not been able to determine whether a nonaggressive story, told to a TAT picture, reflects low intensity of aggressive needs or intense needs coupled with determined self-control. It would seem that such knowledge would be essential to making intelligent decisions about matching workers to patient populations.

The evidence in Table 1 suggests, as other workers have discovered, that all measures of fantasy "aggression" are by no means equivalent. There exists a "hidden" quality in the aggression expressed in the TAT when compared with that revealed in the Scrambled Sentences test. Hostile thoughts, which fit naturally in the flow of events of a story suggested by a TAT card, evoke a much greater self-consciousness when they appear in the form of a sentence which the subject has constructed to read, "I'll shoot you." The high visibility of such responses seems to act as an inhibitor which moves the subject consciously to alter the degree of hostility he allows to make an appearance. This phenomenon is probably most characteristic of somewhat sophisticated subject groups such as our own. In part, the decision to combine the TAT and Scrambled Sentences test into a single measure of aggression was prompted by the reports of several subjects that they changed the direction of their responses after constructing "too many" aggressive sentences. The combined measure held promise of providing some compensation for the subjects' insertion of conscious control over his responses. This measure proved to be an effective device for predicting to the counselor Sociometric Choice but ineffective in foretelling the counselor's ability to resolve conflict and manage a child's open anger.

We found (Table 3) that only one measure (the TAT) was able to predict later behavior in a hostile situation. As is so frequently reported in the literature, all tests reputed to measure "aggression" measure something —although apparently not the same thing. In part, this occurs as a result of the lack of refinement of the various tests, but our logic becomes circular if we allow the blame to rest there. Perhaps it is unreasonable to expect any simple device adequately to account for the multitude of dimensions implicit in an abstraction such as "aggression." The complexity of a typical expression of hostility often involves most of what we know of human behavior and a great deal that we do not. It should not surprise us, then, that tests assessing a broad spectrum of potential response are likely to have some predictive value. In many respects the TAT, drawing as it does from a considerable range of the individual's fantasy life, satisfies this requirement, yet leaves much to be desired as a dependable tool.

A part of this study provides experimental confirmation of what every clinical worker knows. In the test tube of the working situation one's reputation for effectiveness and skill becomes a fundamental aspect of his relations with fellow workers. In the highly visible setting of a therapeutic camp the "office hour" is exposed to scrutiny, and the judgments of peers and supervisors are soon in close agreement. The significance of the relationships between these judgments of clinical adequacy and the aggressive fantasy of the subjects can in part be accounted for by this "open" quality of the camp setting. The frequency of aggressive outbursts by the children presented a host of opportunities to make accurate judgments of the adequacy of adult aggressive control. In Table 2 it is apparent that the accuracy of counselor judgment of the clinical skill of other counselors can be verified by using independent observations of the handling of crisis situations. One's reputation as an effective clinical worker is not always judged with such ease or accuracy since much of therapeutic work is of a highly private nature and leaves only vague inference as a basis for judging. It seems clear that if co-worker judgments are to be used as a means of assessing ability, a significant portion of the process of clinical work must be made visible to others. Too frequently this is not the case; yet we rely heavily on peer opinions in making our judgments.

In such circumstances, the discrete meaning in any individual worker of aggressive fantasy was less important than some estimate of its gross amount. Most clinical settings demand greater refinement of prediction and might profitably make use of the leads provided by this study.

Summary

Our concern with the role of adult aggression in the management of disturbed children arose from the need to appraise the probable success of potential child care workers. From clinical observations it was clear that

the mixing of aggressive children and hostile adults inevitably produced a flood of therapeutic crises which tended to scatter treatment plans into wild disarray. In contrast, the successful management of hostile situations seemed regularly to be characteristic of those adults whose own aggressive needs were the least pressing or under the best control. To test this prediction, two measures of aggression (a Scrambled Sentences Arrangement Test and spontaneous stories in response to four TAT cards) were administered to a group of counselors prior to their summer experience in a therapeutic camp for disturbed and delinquent children. Comparisons were made between the degree of aggressive response to these tests and ratings, by their fellow counselors, of success in managing hostile situations. An additional measure of counselor skill was gained through observation by experienced observers of 10 crisis incidents for each of a group of volunteer counselors. These detailed observations formed the basis for rating both counselor effectiveness in resolving conflict and the degree to which the children were able to accept and "live with" these resolutions of their disputes.

Of the two test measures used to assess the degree of hostility present in the fantasy productions of our sample of child care workers, it is clear that successful prediction to aggressive behavior occurs more readily using the TAT measure. Significant relationships exist between the degree of hostility revealed in spontaneous TAT stories, child care workers' ratings of one another's clinical skill, and experienced observers' ratings of the subject's handling of aggressive situations. A positive and significant relationship was discovered between expert and peer judgments of an individual's clinical ability. The refinement of available testing techniques and further theoretical clarification of the relation of fantasy to behavior is needed to increase the accuracy with which we can predict worker behavior in an aggressive situation.

References

1. Bach, G. R. Young children's play fantasies. *Psychol. Monogr.*, 1945, 59, No. 2 (Whole No. 272).
2. Eriksen, C. W. Some implications for TAT interpretations arising from needs and perception experiments. *J. Pers.*, 1951, 19, 281-288.
3. Feshback, S. The drive-reducing function of fantasy behavior. *J. abnorm. soc. Psychol.*, 1955, 50, 3-11.
4. Holzberg, J. D., Bursten, B., & Santiccioli, A. The reporting of aggression as indication of aggressive tension. *J. abnorm. soc. Psychol.*, 1955, 50, 12-18.
5. Kagan, J. The measurement of overt aggression from fantasy. *J. abnorm. soc. Psychol.*, 1956, 52, 390-393.

6. KORNER, A. F. *Some aspects of hostility in young children.* New York: Grune & Stratton, 1949.

7. MOOD, A. M. *Introduction to the theory of statistics.* New York: McGraw-Hill, 1950.

8. MORSE, W. C. The development of a mental hygiene milieu in a camp program for disturbed boys. *Amer. J. Orthopsychiat.,* 1953, 23, 826-832.

9. MURRAY, H. A. *Thematic Apperception Test: manual.* Cambridge, Mass.: Harvard Univer. Press, 1943.

10. MUSSEN, P. H., & NAYLOR, H. K. The relationship between overt and fantasy aggression. *J. abnorm. soc. Psychol.,* 1954, 49, 235-240.

11. SANFORD, R. N., ADKINS, MARGARET M., MILLER, R. B., COBB, E. A., & OTHERS. Physique, personality and scholarship: a cooperative study of school children. *Monogr. Soc. Res. Child Develpm.,* 1943, 8, No. 1.

12. SEARS, R. R. Relation of fantasy aggression to interpersonal aggression. *Child Develpm.,* 1950, 21, 5-6.

13. WATSON, R. E., PRITZKER, L., & MADISON, P. Hostility in neurotics and normals. *J. abnorm. soc. Psychol.,* 1955, 50, 36-40.

A SELF-CONCEPT SCALE FOR CHILDREN AND ITS RELATIONSHIP TO THE CHILDREN'S FORM OF THE MANIFEST ANXIETY SCALE[1]

LEWIS P. LIPSITT[2]

Iowa Child Welfare Research Station

The kinds of verbalizations that people "make to themselves," perhaps particularly about themselves, have long been of interest to personality theorists such as Murray, Rogers, Sullivan, and Allport. Such self-talk, as verbal behavior, typically becomes important in personality and clinical appraisals. Self-evaluative behavior, in fact, seems to be involved in the operational definitions of certain personality concepts as ego, superego, ego-ideal, self-concept, ideal-self, and the like. Thus, an individual who verbalizes inadequacy and inferiority is said to have weak ego-strength or a low self-concept. Moreover, persons' verbal representations of themselves are often compared with their concepts of an "ideal self" on the not un-

---

[1] A portion of this paper was read at the 1957 meetings of the Midwestern Psychological Association. The writer is indebted to Mr. M. M. Schell, Superintendent, and Principals Elsie E. Glassburner and Edward L. Hutchcroft of the Washington, Iowa, Public Schools for their cooperation in providing subjects and facilities. The writer is appreciative, also, of critical readings of an early manuscript by Drs. Ruth B. Holton and Alfred Castaneda. Thanks are due Mr. Andy E. Rogers who aided in the collection and processing of data.

[2] Now of Department of Psychology, Brown University.

common assumption that in the conflicted or anxious individual there exists a large discrepancy between his ideal and his conception of himself.

There has arisen, however, some evidence contrary to the assumption of a negative relationship between anxiety level and verbal self-estimations. Taylor and Combs (18) presented children with written statements which the authors considered true of all children but which would require self-effacement to admit, finding that those who tended to admit to these statements were those with the highest adjustment scores as measured by the California Test of Personality. Zimmer (19) was unable to find a relationship between a self-acceptance measure and independent measures of conflict derived from a word association task.

Researchers interested in self-concept variables typically utilize discrepancy scores (1, 2, 5, 6, 7, 8, 10, 13, 15, 19) deriving from two test measures, the first involving *S*'s rating himself ("as he really thinks he is") with regard to trait-descriptive adjectives and the second requiring *S* to respond to the same adjectives indicating "how he would like to be." The discrepancy score may consist of the total difference between the two measures, and this discrepancy is considered as reflecting degree of dissatisfaction with oneself.

The present study employs both self-concept and discrepancy measures in fourth, fifth, and sixth grade children and is concerned primarily with (a) the comparative reliabilities of the discrepancy scores and the self-concept scores, and (b) the relationship of each of these two measures to scores on the children's form of the manifest anxiety scale (3, 4, 11, 14).

The self-concept scale contained the following 22 trait-descriptive adjectives, presented here in the order used in the scale: friendly, happy, kind, brave, honest, likable, trusted, good, proud, *lazy,* loyal, co-operative, cheerful, thoughtful, popular, courteous, *jealous,* obedient, polite, *bashful,* clean, helpful. Each of these adjectives was prefaced by the phrase "I am . . ." and was followed by a five-point rating scale. Nineteen were considered as positive or socially desirable attributes, while three were considered negative (italicized above). The rating categories, scored from 1 to 5, were entitled *not at all, not very often, some of the time, most of the time,* and *all of the time.* A score of 1 was received on an item if *S* checked the first category, a score of 5 if the last category were checked, except in the case of the three negative adjectives which were scored in inverse fashion.[3] A score on the self-concept scale was obtained for each *S* by summing the ratings ascribed to himself on each item. Lower scores were presumed to reflect degree of self-disparagement.

The ideal-self scale contained the same adjectives but here each was prefaced by "I would like to be . . . ." Again, *S* rated himself on each item on similar scales. The discrepancy scores employed in the present

[3] The negative words may provide a detection device for nonreaders and indiscriminate checkers. In the present study, however, no *S*s were eliminated on such a basis.

study consisted of the simple subtraction of the total self-concept score from the total ideal-self score.[4]

Both scales were administered, with the children's manifest anxiety scale (CMAS), to approximately 300 fourth, fifth, and sixth graders on the same day, and again two weeks later. Administration was on a group basis in the classrooms, and *Ss* were permitted as much time as needed to complete all the scales. The CMAS was administered first, the self-concept scale next, and the ideal-self scale last.

Reliability of the Children's Form of the Manifest Anxiety Scale

A report of the children's form of the manifest anxiety scale (CMAS), published in this journal (3), contained means, standard deviations, and one-week reliability coefficients for a city population of fourth, fifth, and sixth graders. The present study provides additional information concerning the retest reliability of the CMAS for a two-week interval and a different population of same-aged children, and provides some indication of the generality of the originally reported mean scores for these age groups.

Table 1

ANXIETY SCALE MEANS AND STANDARD DEVIATIONS BY GRADE AND BY SEX

	Fourth Grade			Fifth Grade			Sixth Grade		
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
Boys	47	13.13	7.38	50	15.16	7.96	41	14.51	8.12
Girls	62	15.61	8.83	61	15.26	7.44	37	16.57	8.16

Table 1 contains the anxiety means and standard deviations for each grade and sex separately. These statistics may be compared with those describing the original population (3, p. 320). All of the present means fall somewhat below those previously reported. Significant differences between the present and previous means are found for fourth grade girls only ($t = 2.34$, $p = .05$).[5]

[4] Seven *Ss* who rated themselves higher on the self-concept than the ideal-self scale were eliminated.

[5] Despite the general statistical similarity of these two sets of group means, both of which are derived from presumably normal populations, this one disparity suggests caution in the interpretation of differences found between different test populations (e.g., normal school children vs. clinical samples). That is, significant differences should probably not be taken as necessarily reflecting the presence of population abnormality. It is suggested, furthermore, that the obtained mean for fourth grade girls in the original population may be an overestimate for a population of normal fourth grade girls.

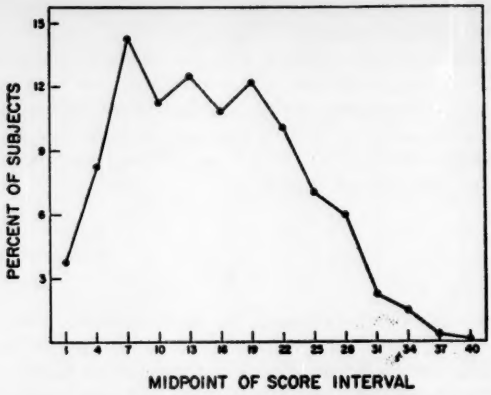


FIGURE 1—Frequency polygon showing percentage of 298 fourth, fifth, and sixth grade children receiving the indicated scores on the children's form of the manifest anxiety scale.

Figure 1 presents a frequency polygon comparable to that of the original group (3, p. 320). The present group provides a somewhat flatter distribution with a tendency for the 20, 50, and 80 percentile positions to be lower by about two score-units (7.18, 14.41, and 22.54, respectively). The present coefficients (Table 2) are generally comparable to those obtained previously (3, p. 321). Each is reliable beyond the .001 level. It should be noted, however, that, whereas the first study provided its lowest reliability coefficient for the fourth grade girls relative to the remaining groups, the present study yielded its highest coefficients for that group.

The lie scale means and standard deviations for the present group are reported in Table 3, and may be compared with those of the original population (3, p. 321). All of the present means are greater than those previously reported. None of the correlations between anxiety and lie scores for each grade and sex separately reaches significance, ranging from —.20 to .16 in

TABLE 2

ANXIETY SCALE TEST-RETEST CORRELATIONS BY GRADE AND BY SEX

	Fourth Grade		Fifth Grade		Sixth Grade	
	N	*r*	*N*	*r*	*N*	*r*
Boys	44	.89	48	.87	41	.91
Girls	57	.93	52	.89	35	.83

the present study. A chi square test indicates that these r's constitute only chance deviations from a common population value; pooling these r's for all grade-sex combinations yields a composite r of $-.114$, which falls short of significance at the .05 level. Thus, as in the previous study, the tendency to falsify responses on the anxiety scale may result in either a high or low anxiety score.

TABLE 3

LIE SCALE MEANS AND STANDARD DEVIATIONS BY GRADE AND BY SEX

	Fourth Grade			Fifth Grade			Sixth Grade		
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
Boys	47	4.57	1.95	50	3.84	2.26	41	3.15	2.18
Girls	62	5.58	1.77	61	4.16	2.48	37	3.95	2.63

The Self-Concept and Ideal-Self Scales

Table 4 presents the means and standard deviations for the self-concept scale for each grade and sex separately. An analysis of variance was performed and neither the main effects of grade or sex nor the interaction of these two variables was found significant. The six means were pooled to obtain an estimate of a mean score common to the combined population of fourth, fifth, and sixth graders. This mean is 86.75 with a standard deviation of 8.18.

TABLE 4

SELF-CONCEPT SCALE MEANS AND STANDARD DEVIATIONS BY GRADE AND BY SEX

	Fourth Grade			Fifth Grade			Sixth Grade		
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
Boys	47	85.81	9.75	50	86.24	8.25	41	87.17	6.85
Girls	62	87.39	10.07	61	86.74	8.27	37	87.11	8.76

Figure 2 represents a frequency polygon showing the percentage of *Ss* receiving the indicated scores on the self-concept and ideal-self scales. The distribution of scores for the self-concept scale is essentially normal, as contrasted with the distribution for the ideal-self scale, where a J-curve results with most *Ss* receiving high scores. Thus, most *Ss* rate themselves as being highly desirous of attaining the positive attributes of the scales.

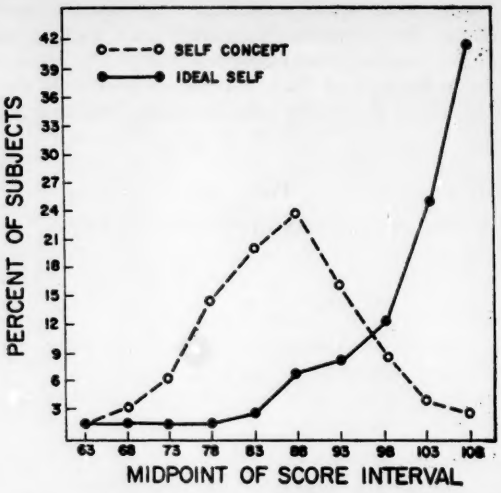


FIGURE 2—Frequency polygon showing percentage of 298 fourth, fifth, and sixth grade children receiving the indicated scores on the self-concept and ideal-self scales.

Our interest here is in the respective reliabilities of the *self-concept* and *discrepancy* scores. Table 5 depicts the two-week test-retest reliability coefficients for the self-concept measure alone. All of these correlations are significant beyond the .001 level.

Also in Table 5 are the comparable retest correlations for the discrepancy measure. These coefficients are all somewhat below those for the self-

TABLE 5

TEST-RETEST CORRELATIONS FOR SELF-CONCEPT SCALE AND DISCREPANCY SCORES BY GRADE AND BY SEX

	Fourth Grade		Fifth Grade		Sixth Grade	
	N	*r*	*N*	*r*	*N*	*r*
Self-Concept Scale						
Boys	44	.73***	46	.80***	40	.84***
Girls	56	.78***	49	.91***	35	.80***
Discrepancy Scores						
Boys	28	.57**	44	.72***	38	.56***
Girls	32	.66***	48	.68***	34	.51**

** Significant at .01 level. *** Significant at .001 level.

concept measure.[6] Four of six attain significance, nevertheless, at the .001 level, and two at the .01 level.

Relationships of Self-Concept and Ideal-Self Measures with Scores on the Children's Manifest Anxiety Scale

Table 6 reveals a negative relationship between the self-concept scores and anxiety scores as measured by the CMAS, for each of the grades and sexes separately. In all cases, these correlations are significant, five at the .01 and one at the .05 level. The magnitude of these correlations may be compared with those relating anxiety to the discrepancy scores, only four of which attain significance, two at the .01 and two at the .05 level. It would appear, then, that the self-concept score by itself yields a measure that is more reliable than the discrepancy score and that is more highly related to the CMAS score.

Table 6

CORRELATIONS BETWEEN ANXIETY SCALE AND SELF-CONCEPT SCALE AND DISCREPANCY SCORES BY GRADE AND BY SEX

	Fourth Grade		Fifth Grade		Sixth Grade	
	N	*r*	*N*	*r*	*N*	*r*
Anxiety Scale and Self-Concept Scale						
Boys	47	−.53**	50	−.40**	41	−.34*
Girls	62	−.63**	61	−.40**	37	−.58**
Anxiety Scale and Discrepancy Scores						
Boys	43	.31*	48	.24	40	.20
Girls	55	.51**	60	.39**	36	.39*

* Significant at .05 level. ** Significant at .01 level.

Conclusions and Discussion

The conclusions that may be drawn from these data are: (a) both the self-concept and discrepancy measures, as presented here, provided reliable estimates of an individual difference variable; (b) of these two types of measures, the self-concept score provided a somewhat more reliable measure than the discrepancy score; (c) for the fourth, fifth, and sixth grade *Ss*, there were no reliable grade or sex differences in mean self-concept scores; (d) for all six grade-sex combinations, the self-concept measure correlated significantly with performance on the CMAS, while comparable correla-

[6] The loss of *Ss* for the fourth grade in Table 6 was due to the impossibility of administering the ideal-self scale to one of the classes on the second day of testing.

tions for the discrepancy scores were somewhat less and in some cases not reliable.[7]

The tendency toward self-disparagement, as reflected in responses to the self-concept scale, was found to be significantly related to another independent measure of the degree to which a person admits to possessing symptoms commonly held to reflect anxiety.[8] Several theorists (9, 16) have ascribed drive properties to anxiety; thus, theoretical predictions concerning the role of drive in learning may be applicable here as well.

The tendency toward self-disparagement may be conceived as a phenomenon producing behavioral effects consonant with its status as a drive variable as follows: The conditioning principles presumed to underlie the acquisition of secondary drives assume that neutral stimuli preceding reinforced responses may come through training to elicit those responses themselves in the absence of the original reinforcement. Thus, if a verbal stimulus ("You are naughty") is presented in conjunction with a noxious stimulus (physical punishment) which unconditionally elicits an emotional response (pain), that verbal stimulus may acquire the capacity to elicit (some part of) the emotional response independent of the original noxious stimulus. It is quite conceivable, then, that the conditions (behavior of the child) which originally resulted in the noxious stimulation (punishment) may come to elicit the verbal response and its stimulus counterpart ("I am naughty") which in turn mediates the emotional response. The tendency toward self-disparagement becomes, then, an antecedent condition for general emotional responsiveness (anxiety) or drive.

Summary

The children's form of the manifest anxiety scale and self-concept and ideal-self scales were administered twice to approximately 300 fourth, fifth, and sixth graders at a two-week interval. A discrepancy score, or measure of self-disparagement, was obtained by subtracting *Ss*' self-concept ratings from their ideal-self ratings. It was found that the self-concept measure taken by itself was more reliable than the discrepancy measure, and that

[7] The lower retest reliabilities of the discrepancy measure may be attributed to the fact that difference scores always involve additional sources of variability to those involved in a single test. The coefficient for the discrepancy measure is a function of the variability of the two tests, since the discrepancy score is derived from the use of both the self-concept and ideal-self scores. McNemar (12, p. 154) points out that the standard error of such difference scores must be lower than for one test alone, and that this tends to depress the reliability estimates of difference scores.

[8] Some may consider this relationship as indicative of the presence of a response set or test-taking attitude which mediates or influences scores on both scales. The view taken here is that whenever two test measures correlate, it may be said that a response set is responsible in that there exists a source of variance common to both measures. When test measures are correlated, they are measuring something in common, response set or whatever.

the self-concept measure was more highly related to CMAS score than was the discrepancy score. Significant correlations were obtained for all grades and sex combinations between CMAS and self-concept scores, with high anxious *Ss* producing low self-concept (or high self-disparagement) ratings. The two-week reliability of the CMAS was essentially the same as that for the original population.

References

1. Bills, R. E., Vance, E. L., & McLean, D. S. An index of adjustment and values. *J. consult. Psychol.*, 1951, 15, 257-261.
2. Block, J., & Thomas, H. Is satisfaction with self a measure of adjustment? *J. abnorm. soc. Psychol.*, 1955, 51, 254-259.
3. Castaneda, A., McCandless, B. R., & Palermo, D. S. The children's form of the manifest anxiety scale. *Child Develpm.*, 1956, 27, 317-326.
4. Castaneda, A., Palermo, D. S., & McCandless, B. R. Complex learning and performance as a function of anxiety in children and task difficulty. *Child Develpm.*, 1956, 27, 327-332.
5. Chordokoff, B. Adjustment and the discrepancy between the perceived and ideal self. *J. clin. Psychol.*, 1954, 10, 266-268.
6. Cowen, E. L., Heilizer, F., & Axelrod, H. S. Self-concept conflict indicators and learning. *J. abnorm. soc. Psychol.*, 1955, 51, 242-245.
7. Friedman, I. Phenomenal, ideal, and projected conceptions of self. *J. abnorm. soc. Psychol.*, 1955, 51, 611-615.
8. Hanlon, T. E., Hofstaetter, P. R., & O'Connor, J. P. Congruence of self and ideal self in relation to personality adjustment. *J. consult. Psychol.*, 1954, 18, 215-218.
9. Hull, C. L. *Principles of behavior*. New York: D. Appleton-Century, 1943.
10. Levy, L. H. The meaning and generality of perceived actual-ideal discrepancies. *J. consult. Psychol.*, 1956, 20, 396-398.
11. McCandless, B. R., & Castaneda, A. Anxiety in children, school achievement, and intelligence. *Child Develpm.*, 1956, 27, 379-382.
12. McNemar, Q. *Psychological statistics*. New York: Wiley, 1955.
13. Miller, K. S., & Worchel, P. The effects of need-achievement and self-ideal discrepancy on performance under stress. *J. Pers.*, 1956, 25, 176-190.
14. Palermo, D. S., Castaneda, A., & McCandless, B. R. The relationship of anxiety in children to performance in a complex learning task. *Child Develpm.*, 1956, 27, 333-337.

15. Roberts, G. E. A study of the validity of the index of adjustment and values. *J. consult. Psychol.*, 1952, 16, 302-304.

16. Spence, K. W. Current interpretations of learning data and some recent developments in stimulus-response theory. In *Kentucky Symposium: Learning Theory, Personality Theory, and Clinical Research*. New York: Wiley, 1954.

17. Spence, K. W. *Behavior theory and conditioning*. New Haven: Yale Univer. Press, 1956.

18. Taylor, C., & Combs, A. W. Self-acceptance and adjustment. *J. consult. Psychol.*, 1952, 16, 89-91.

19. Zimmer, H. Self-acceptance and its relation to conflict. *J. consult. Psychol.*, 1954, 18, 447-449.

SIZE JUDGMENT AT DIFFERENT DISTANCES AS A FUNCTION OF AGE LEVEL

WALTER COHEN, AARON HERSHKOWITZ, *and*
MARJORIE CHODACK
University of Buffalo

Although many investigators and theorists have dealt with the problem of perceptual constancy, an adequate understanding of the phenomenon is lacking. Much of the controversy has been concerned with the traditional problem of nativism vs. empiricism. Whether or not this issue can be resolved, the development of constancy would seem to merit study in and of itself.

Studies of constancy as a function of age have yielded inconsistent data. Size constancy represents a clear case in point. Various investigators (1, 2, 9) have found that size constancy increases with age while others have presented evidence that there is no such relationship (3, 4). Variations in experimental methodology and measurement techniques have made direct confirmation of theory and comparison of data impossible. The experiments differ considerably in the shape, size, and distance of stimuli, types of comparison made, instructions, and range of ages tested. Any one of these variables might well affect the outcome of a developmental study. In addition, very small samples and inadequate statistical analysis characterize most of the research in this field.

There appears to be an implicit assumption that size constancy is a simple, unitary phenomenon which develops, or does not develop, with age, and there is thus a tendency to ignore the specific experimental conditions under which constancy is measured. The present experiment is designed to investigate some of the variables which might account for the seemingly contradictory data in the literature and to clarify the issue of constancy as a unitary process.

One of the variables that seems most significant is that of the distance of the objects. Lambercier and Piaget (5) found size constancy differences between children and adults for "far" objects. Young children (5 to 7) tended toward underconstancy; older children (10 to 11) showed near perfect constancy, while adults (17 and above) showed definite overcon-

CHILD DEVELOPMENT, Vol. 29, No. 4 (December, 1958)

stancy. No differences were found between children and adults for "near" distances, and Piaget contends that under these conditions constancy is complete by the end of the first year. The terms "near" and "far" appear to refer mainly to the distance between the observer and the most distant object. More recently, Zeigler and Leibowitz (9) found that 7- and 8-year-olds show good constancy at "near" distances (10 ft.) while adults show overconstancy at this distance. Both groups display underconstancy at greater distances, but the adults are more accurate in their judgments. In both of these investigations, the variable of distance is ambiguous since both the distance of objects from the observer, as well as the distance between the objects themselves, are simultaneously varied. The present study was designed to investigate the effects of differences in depth between the test objects themselves.

Werner (7) has stressed the diffuse quality of the child's perception. However, this aspect has been neglected in previous studies of size estimation development. Usually the child is restricted to a single estimation of equivalence and only the PSE is computed. The present investigators used the Method of Limits, which permits the *O* to give a number of responses of "same" to the standard during a single series (8). It was felt that a measure of the range of equivalence, i.e., the "interval of uncertainty" (IU), might reveal aspects of size judgments which are relevant for the problem of perceptual development.

Method

Apparatus

The experiment was conducted in a well-illuminated room, 9 by 30 ft., in which the walls and windows were entirely covered with white paper. *O* sat on an adjustable chair. His chin was placed on a headrest 100 cm. from the floor. Three black cardboard squares, 5, 6, and 7 cm. each, served as standard stimuli. Only one of these was present at each trial. Two desks, each 73 cm. high, were placed at 2 and 6 m. from the *O* and were visible at all times. A standard was clamped to one of the two desks so that it was at a height of 129 cm. from the floor. The standard was always 30 cm. to the left of *O*'s line of sight.

Thirteen black cardboard comparison stimuli were mounted on white dowels which were in turn mounted on a rotating drum. The drum could be rotated so that only one comparison stimulus was exposed at a time at a height of 169 cm. from the floor. The sizes of the comparison stimuli ranged in ½ cm. steps from 3 cm. to 9 cm. and were arranged consecutively on the drum. A white cloth, which obscured both the *E* and the comparison wheel from *O*'s view, was stretched across the width of the room directly in front of the drum.

To summarize, each *O* was confronted with a single comparison stimulus (at a distance of 8 m. and directly in front of *O*) and a single standard

(at a distance of either 2 or 6 m. and slightly to the left of *O*). Stimuli were seen against a white background. The only other objects in the field were two desks, placed 2 and 6 m., respectively, in front of *O*.

Subjects

Forty-two *O*s were randomly selected from various classes of the Barker, N.Y., school system. There were 12 *O*s each in the 7-, 12-, and 17-year-old groups, and 6 *O*s in the 5-year-old group. An equal number of boys and girls was selected for each age group. Only one *O*, a 5-year-old, was not testable and was replaced.

Procedure

Each *O* was seated at an appropriate height so that he was comfortable when using the head rest. He was instructed to tell whether the card above the white cloth looked larger than, smaller than, or the same size as the card on the desk. The instructions were repeated and the objects were pointed to when necessary. The tester was a 22-year-old woman who was previously known to most of the *O*s.

An ascending and descending series of judgments were obtained from *O* for each of the 3 standards at both distances (2 and 6 m.). The order of presentation was counterbalanced for distance and size of the standard as well as for the order of ascending and descending series. The size of the comparison card was also systematically varied for each series.

Results

The mean judgments as well as the average constant error (CE) and average IU for each age level are presented in Table 1. A separate analysis

Table 1

MEAN PSE, CE, AND IU FOR EACH AGE LEVEL

Distance of Standard	*Age*	PSE *Size of Standard* 5 cm.	6 cm.	7 cm.	*Over-all Average for All Standards* CE	IU
2 m.	17	5.40	6.44	7.16	.33	.72
2 m.	12	5.82	6.81	7.51	.71	.84
2 m.	7	5.71	6.93	7.67	.77	.95
2 m.	5	6.10	6.94	7.94	.99	1.42
6 m.	17	5.29	6.36	7.22	.29	.48
6 m.	12	5.39	6.52	7.27	.39	.60
6 m.	7	5.34	6.52	7.25	.37	.86
6 m.	5	5.44	6.42	7.15	.33	1.16

of variance was computed for each size standard. Although a significant F ($p < .01$) was obtained for distances (2 vs. 6 m.) in each of the three analyses, the F for age levels was significant ($p < .05$) only for the 5 cm. standard. The judgments of the standards at 6 m. were more accurate than those at 2 m. In addition, there were no consistent differences between the various age levels at the 6 m. distance.

In order to evaluate the over-all performance for all three standards at the 2 m. distance, a separate analysis was performed. The CE, which is obtained by subtracting the PSE from the size of the standard, was used in this analysis (8). It should be noted that differences in average CE are identical in numerical value to differences in average PSE for each age level. However, the use of the CE makes comparison of accuracy with the different size standards possible. This analysis appears in Table 2.

TABLE 2

ANALYSIS OF VARIANCE OF CE WITH THE STANDARDS AT 2 M.

Source	*df*	*M.S.*	*F*	*p*
Size (S)	2	.716	1.81	>.10
Age (A)	3	2.140	5.42	.005
S × A	6	.088	.22	
Individuals	114	.394		
Total	125			

A significant F ($p < .005$) was obtained for age levels. The magnitude and statistical significance of the differences between the mean CE for the four age levels appear in Table 3. The 17-year-olds were more accurate than all the other age groups ($p < .01$) when the standard was at 2 m. None of the younger age groups differed significantly from one another. There was also no significant difference in the accuracy of judgment of the three standard objects.

TABLE 3

DIFFERENCES IN OVER-ALL CE AMONG THE FOUR AGE LEVELS

	STANDARD AT 2 M.			STANDARD AT 6 M.		
Age	*12*	*7*	*5*	*12*	*7*	*5*
17	.38*	.44*	.66*	.10	.08	.04
12		.06	.28		—.02	—.06
7			.22			—.04

* $p < .01$.

In the analysis of the IU, nonparametric procedures were utilized because the distributions were skewed. By means of the sign test (6), a significant difference ($p < .05$) was obtained between the mean IU at 2 m. and that at 6 m. Thus, there was a reduction of equivalence at 6 m. An analysis of the differences between age levels, using White's extension of the method of Wilcoxon (6), indicated that both the 17- and 12-year-olds had a significantly lower IU than the 5-year-olds at 2 m. ($p < .05$) and also at 6 m. ($p < .01$). In addition, the 17-year-olds differed significantly ($p < .01$) from the 7-year-olds at 6 m.

Although no differences were found between the various age groups in accuracy, i.e., in size constancy, at the 6 m. distance, the groups differed considerably in IU at this distance. The implications of this finding as well as the other results will be discussed in the following section.

Discussion

The above results clearly show that young adults have better size constancy than children when the distance between standard and comparison objects is relatively large and that these differences disappear when the distance between the objects is reduced. In the light of these findings, it is possible to reinterpret some of the contradictory data in the size constancy literature. Above all, it is now clear that the conception of size constancy as a unitary phenomenon is unwarranted.

Upon re-examination of the evidence on both sides of the so-called nativism-empiricism controversy, it becomes apparent that supporters of the empiricist position (1, 2, 9) have tended to emphasize those results within an experiment which demonstrated superiority of the adults over the children. At the same time, they tended to neglect any indications of similarity between the age groups. The nativists did exactly the reverse (3, 4).

One of the factors which seems to account for variations in the results of previous investigations is that of the distance between the two stimulus objects. For example, Beryl's data (1), which have often been used to support a learning approach, clearly indicate that the differences in size constancy between age levels decrease as the distances between test objects decrease. When this distance is only 2 m., there are no differences in judgment after 3 years of age. Zeigler and Leibowitz (9) also tend to stress the differences between adults and children when the distances between test objects are large, i.e., 25 ft. and over, but ignore the equal size constancy of the children and adults when the distance between the test objects is small, i.e., 5 ft.

The nativists often cite the investigations by Burzlaff (3) in which near perfect size constancy was found at all age levels. This performance has often been attributed to the method he utilized, which consisted of the

simultaneous presentation of the entire series of comparison objects. However, in Burzlaff's experiments, the stimulus objects were separated by a distance of only 3 m. and it is conceivable that it is this small distance between test objects which accounts for the similarity between age groups.

Both Koffka (4) and Brunswik (2) have suggested that the reduction in difficulty of a perceptual judgment will tend to reduce greatly the differences in accuracy between age groups. It seems reasonable to assume that the closer together two stimuli are the less difficult the task of size estimation will be. This assumption is supported by the results of the present study: (a) a reduction in the differences in CE between age levels; (b) an actual reduction of the CE; and (c) a reduction in the IU.

It should be pointed out that although the 17-year-olds in the present study were more accurate in their judgments, there were no reliable differences among the three younger age groups. Although a cursory examination of the data might lead one to interpret the performance of the younger groups as an indication of a developmental trend, such an interpretation would be erroneous. The differences between the means do not approach significance, and the median CE for each of the three younger groups is identical. It would seem, then, that when the standard is 2 m. from *O* and the comparison stimulus is at 8 m., there is very little change in size constancy from age 5 to age 12, but an increase in size constancy does occur between the ages of 12 and 17.

Even when the perceptual judgment is relatively simple, as in the 6 m. condition, and there are no differences in CE between age levels, differences in the IU are evident. The 17- and 12-year-olds both show a significantly smaller IU than the 5-year-olds, while the 7-year-olds occupy an intermediate position. Inasmuch as a high IU indicates an inability to differentiate a large number of different sized objects, the performance of the younger children can be interpreted as a reflection of their greater diffuseness in perception. It should be emphasized that, although achievement in terms of accuracy may be similar for different age groups, differences in other aspects of performance may still exist.

Summary

Judgments between a comparison object placed 8 m. from *O* and a standard placed at either 2 m. or 6 m. from *O* were obtained from 42 boys and girls. The 17-year-olds were significantly more accurate in size estimation than the younger groups (5-, 7-, and 12-year-olds) when the standard was at 2 m. from *O* but not when it was at 6 m. The 17- and 12-year-olds had a lower IU than the 5-year-olds at both distances. These results suggest that the distance between test objects is an important variable in understanding the contradictory data obtained in developmental studies of size constancy. One of the major implications of the study is that size constancy should not be treated as a unitary phenomenon.

References

1. Beryl, F. Über die Grössenauffassung bei Kindern. *Z. Psychol.*, 1926, 100, 344-371.
2. Brunswik, E. *Perception and the representative design of psychological experiments.* Berkeley & Los Angeles: Univer. of California Press, 1956.
3. Burzlaff, W. Methodologische Beitrage zum Problem der Farbenkonstanz. *Z. Psychol.*, 1931, 119, 117-235.
4. Koffka, K. *Principles of Gestalt psychology.* New York: Harcourt Brace, 1935.
5. Piaget, J. *Psychology of intelligence.* New York: Harcourt Brace, 1950.
6. Snedecor, G. W. *Statistical methods.* (5th Ed.) Ames, Iowa: Iowa State College Press, 1956.
7. Werner, H. *Comparative psychology of mental development.* New York: Harpers, 1940.
8. Woodworth, R. S., & Schlosberg, H. *Experimental psychology.* (Rev. Ed.) New York: Holt, 1954 .
9. Zeigler, H. P., & Leibowitz, H. Apparent visual size as a function of distance for children and adults. *Amer. J. Psychol.*, 1957, 70, 106-109.

DEVELOPMENTAL, PSYCHOSOCIAL, AND EDUCATIONAL FACTORS IN CHILDREN WITH NONORGANIC ARTICULATION PROBLEMS[1]

RUTH FITZSIMONS

Warwick School Department
Warwick, Rhode Island

Nonorganic, articulatory problems, involving sound substitutions, omissions, distortions and additions are the most common of all speech disorders. One research study (10) has placed the incidence of articulatory problems at 50 per cent of all speech difficulties which are represented among school children. Investigation in this area is justified by the numerical supremacy of the problem (13) and by the comparative paucity of available research (9, 13).

An articulation problem is a manifestation that the child's speech development is impaired and that his language progression is protracted. The primary problem investigated in this study may be stated as follows: Do children who have nonorganic, articulatory problems differ from children who speak normally in ways other than in the speech disorder syndrome? The rationale of the present investigation was that complete understanding of the child and his articulatory problem comes only by means of an organismic study of the child in the light of his developmental, educational, psychosocial, and speech attributes.

SUBJECTS

The children for this study were drawn from a sample population of 1,477 first grade children in the public school system of Warwick, Rhode Island. The subjects were divided into two groups consisting of 70 children each. The control group was comprised of 70 children whose speech was identified as normal, the experimental group, of 70 children whose speech

[1] This article is based on an Ed.D. dissertation completed under the direction of Dr. Wilbert L. Pronovost, Director of the Boston University Speech and Hearing Center.

handicaps were diagnosed as diffuse, nonorganic, articulatory problems. The group members were screened and found to be free from physical anomalies of the speech and otic mechanisms. On the basis of the results obtained from sweep frequency tests of hearing acuity, only those children whose hearing acuity was found to be within the normal limits of 15 decibels were included in the normal and articulation samples.

The matching criteria for the two groups were CA within four months; IQ, within five points; sex and school locale. The articulation group included 40 males and 30 females, ranging in age from 6 years, 2 months, to 7 years, 10 months (median, 6 years, 7 months). The normal group consisted of 40 males and 30 females who ranged in age from 6 years, 2 months to 7 years, 9 months (median, 6 years, 8 months). For the articulation group, IQ's ranged from 83 to 123 (median, 108), and from 81 to 120 for the normal group (median, 107).

Procedure

Historical case study data were obtained on the 140 subjects by means of a rapport-type, personal interview of the mother by the experimenter at the home of each child. The following data were secured:

1. Child's date of birth and chronological age
2. Type of pregnancy
3. Type of birth
4. Type of labor
5. Child's age at the time of weaning from the bottle to the cup
6. Child's age at the time toilet training was implemented
7. Child's age at the time toilet training was accomplished
8. Age of onset of speech:
    a. word
    b. phrase
9. Age of onset of locomotion (walking)
10. Illnesses which occurred during the period of neonate through the sixth year of life
11. Presence or absence of the following childhood problems as chief complaints at the time of the interview:
    a. destructiveness
    b. eating and food problems
    c. enuresis
    d. fears
    e. finger nail biting
    f. the hurting of pets
    g. jealousy
    h. lying
    i. nervousness
    j. refusal to obey
    k. showing off
    l. shyness
    m. sleeplessness
    n. temper tantrums
    c. thumb sucking

The reading readiness status of each subject was determined by means of the Metropolitan Reading Readiness Test, Form R (5), administered in September, 1953. In February, 1954, each child in the first grade sample population was administered the Kuhlmann-Anderson Group Test of Intelligence (6). At the close of the academic year in May, 1954, the achievement level of each child was determined by the Metropolitan Achievement Tests (4). Final report card grades were also available for analysis. The Vineland Social Maturity Scale (3) was used as a part of the parental interview.

The Children's Apperception Test was administered and the protocols interpreted on a descriptive level, noting the average number of words per protocol and the absence or presence in each protocol of the following dynamisms: orality, hostility directed toward parental figures, perception of parental figures as authoritarian, aggressive phantasies, sibling rivalry, fears and anxieties, significant conflict between autonomy-compliance, and positive or negative outcome. Three judges, one the experimenter, analyzed the protocols of 10 per cent of the subjects, selected by ordering the children in the total sample from highest to lowest IQ and then choosing every tenth child. No significant difference was found among judges in the problem areas identified.

Each child was rated in May, 1953, by his teacher on A Teacher Administered Rating Chart for Talks at the Primary Level (11) with respect to knowledge of topic, organization and content of topic, appearance, voice, and amount of participation. This study was concerned with the total score obtained by each child.

TABLE 1

COMPARISON OF ARTICULATION AND NORMAL SAMPLES ON DEVELOPMENTAL FACTORS

Developmental Factors	*Articulation Sample* ($N=70$)	*Normal Sample* ($N=70$)	*Chi Square*
Abnormal birth	16	7	4.21*
Bottle feeding	68	62	3.87*
Weaning before 6 months of age	28	6	18.80*
Toilet training before 9 months of age	37	24	4.91*
Accomplishment of toilet training before 2 years of age	22	11	4.80*
Occurrence of childhood diseases before or at the age of 3 years	40	20	11.67*
Walking before 12 months of age	18	33	6.94*
Talking (words) before 12 months of age	17	36	10.96*
Talking (phrases) before 18 months of age	17	36	10.96*

* Significant at .05 level.

Null hypotheses pertaining to the discrete data of this study were tested by chi square for correlated samples. The continuous data were subjected to the *t* test for related measures.

Findings

Table 1 reveals that significantly more members of the articulation group experienced abnormal birth conditions, bottle feeding, the imposition of early weaning (prior to 6 months of age), the early occurrence of childhood diseases, early age of implementation and accomplishment of toilet training, and delay in both locomotive and communicative attempts.

Table 2

COMPARISON OF ARTICULATION AND NORMAL SAMPLES ON PRESENCE OF CONDUCT AND HABIT DISORDERS

Conduct and Habit Disorders	*Articulation Sample* (*N*=70)	*Normal Sample* (*N*=70)	*Chi Square*
Destructiveness	18	2	14.93*
Eating and food problems	39	2	47.22*
Enuresis	9	3	3.28
Fears	21	10	5.01*
Finger-nail biting	7	11	1.10
Hurting pets	1	0	.11
Jealousy	6	3	5.08*
Lying	6	1	3.77
Nervousness	53	22	27.60*
Refusal to obey	27	7	15.54*
Showing off	30	6	21.54*
Shyness	31	13	10.74*
Sleeplessness	6	1	3.77
Temper tantrums	35	8	24.47*
Thumb sucking	34	15	11.33*

* Significant at .05 level.

Table 2 indicates that destructiveness, eating and food problems, fears, jealousy, nervousness, refusal to obey, showing off, shyness, temper tantrums, and thumb sucking existed to a significantly greater extent in the children who presented articulation disorders. No significant differences were apparent between the two groups in these psychologic problems: enuresis, finger-nail biting, hurting pets, lying, and sleeplessness.

The incidence of unsatisfactory report card grades in health habits, work habits, and language as well as unsatisfactory scores in reading readiness status and reading grade placement was significantly greater within

TABLE 3

COMPARISON OF ARTICULATION AND NORMAL SAMPLES ON EDUCATIONAL FACTORS

Educational Factors	*Articulation Sample* (*N*=70)	*Normal Sample* (*N*=70)	*Chi Square*
Below average reading readiness status	25	12	6.21*
Unsatisfactory reading grade placement on achievement test	26	12	7.08*
Nonpromotion to grade 2	16	9	2.39
Unsatisfactory report card grade in handwriting	12	7	1.52
Unsatisfactory report card grade in health habits	16	3	10.29*
Unsatisfactory report card grade in language	17	1	14.34*
Unsatisfactory report card grade in reading	17	8	3.94*
Unsatisfactory report card grade in social habits	15	7	3.45
Unsatisfactory report card grade in work habits	19	6	8.22*

* Significant at .05 level.

the speech handicapped sample. No relationship was found between membership in the articulation group and frequency of nonpromotion to grade two, unsatisfactory grades in handwriting or in social habits. (*See* Table 3.)

Table 4 summarizes the results of the analysis of the CAT protocols. The articulation group exceeded to a significant degree the normal group in aggression, fears and anxieties, and perception of parents as authoritarian. The mean number of positive outcomes was significantly greater for the normal children.

TABLE 4

DIFFERENCES BETWEEN THE ARTICULATION AND NORMAL GROUPS ON THE *CAT* AND ON ORAL PARTICIPATION

	$M_a - M_n$	*t*	*p*
CAT Dynamisms			
Aggression	1.30	3.74	.01
Conflict between autonomy-compliance	.50	1.81	*ns*
Fears and anxieties	.30	8.10	.01
Hostility directed toward parental figures	.29	2.00	*ns*
Mean no. of words per protocol	7.26	1.86	*ns*
Orality	.10	.04	*ns*
Parents perceived as authoritarian	1.20	4.26	.01
Positive outcomes	—1.26	3.04	.01
Sibling rivalry	.03	.62	*ns*
Total Oral Participation Score	—6.04	4.17	.01

The mean total oral participation score was significantly higher for the normal group (Table 4). There were no significant differences on the Vineland Social Maturity Scale.

Discussion

The thesis that language problems appear to be psychogenic in etiology has been given empirical support (7). Mowrer (8), considering speech development in the light of psychoanalytic and modern learning theories, emphasized the importance of opportunities for the child to attend to auditory stimuli under conditions which are conducive to pleasurable associations. The child's interactions with the important people in his psychosocial milieu are the origins out of which language learnings develop. Too early and too rigid an imposition of demands for weaning and toilet training are considered to portend subsequent anxiety and personality problems (2, 12). An audit of the anamneses of the articulation sample reveals that this group experienced an early initiation of both weaning and toilet training in addition to the early onset of childhood illnesses. Such a combination of factors may be considered to constitute negative valences in the life experiences of the children who comprised the articulation sample during the decisive period of their language learning. The articulation group tended to be delayed in both their communicative and locomotive attempts, talking and walking at later ages than did the normal sample. The articulation group members were reported to be bottle fed; because the difference between the articulation sample and the normal sample was not great, no conclusions concerning this variable can be drawn. Historical data also revealed that the incidence of abnormal birth conditions was significantly greater in the articulation sample.

The prevalence of tension-reducing symptoms in the articulation subjects is indicated by the significantly greater number of such behavior and conduct disorders as destructiveness, eating and food problems, fears, jealousy, nervousness, refusal to obey, showing off, shyness, temper tantrums, and thumb sucking.

With regard to the inferior status of the children in the articulation group in language, in health habits (which included, among other items, the avoidance of putting things into the mouth), and in oral participation, it may be assumed that the teachers, in their evaluations, were influenced negatively by the presence of articulatory errors in speech. It is interesting to note that each of the factors has an oral component.

The CAT protocols with their references to destruction, injury, and mutilation, their projection of fears and anxieties, their perceptions of parents in authoritarian, demanding roles, and their negative outcomes all suggest a basically negative orientation in these subjects.

The failure to find differences between the normal and articulation subjects in social competence may be explained, in part, by the fact that

the child with the articulation problem finds in the acquisition of social skills and proficiencies, particularly as reinforced by parents, a modality for gaining recognition.

The children who presented diffuse, nonorganic articulation problems differed from the children who spoke normally in ways other than in the speech disorder syndrome. The findings of this study suggest that a relationship may exist between a nonorganic articulation disorder and psychological factors. Two tentative hypotheses may be postulated. A nonorganic speech problem may be satisfying a need for the reduction of internal tension. Such a speech disorder may arise when the mechanism of identification (vital to language learnings) fails to prosper within the child whose early psychosocial milieu is surcharged negatively. Additional research on the relationship between a diffuse nonorganic articulatory disorder and psychological aspects is indicated.

The findings and implications of this study do serve to support Travis' (14) promulgation relative to the limitations of an essentially symptomatic treatment approach in speech therapy. In recognition of the possible relationship between a nonorganic articulation problem and psychologic factors, it appears that an efficacious therapeutic process requires a complete and continuous study of the child, his total needs, his adaptive mechanisms, and his interpersonal relationships as well as a consideration of his articulatory aberrations.

Summary

The purpose of this study was to determine if speech handicapped children differed from normal speaking children in ways other than in the speech disorder syndrome. Data were gathered from a study of two matched groups: a control group of 70 children with normal speech and an experimental group of 70 children with diffuse, nonorganic articulatory problems. Historical case study data were secured. Projective and nonprojective test results were analyzed for both groups. The articulation sample revealed more deviation in the developmental, psychosocial, and educational variables tested than did the normal speaking sample. The major findings were:

1. Members of the articulation group experienced significantly more abnormal birth conditions, bottle feeding, early weaning, the occurrence of childhood diseases during the first three years of life (a consequential period for language learning), the implementation and the accomplishment of toilet training at an early age, as well as delay in both locomotive and communicative attempts.

2. The conduct and habit disorders of destructiveness, eating and food problems, fears, jealousy, nervousness, disobedience, showing off, shyness, temper tantrums, and thumb sucking existed to a significantly greater extent in the speech handicapped sample.

3. Unsatisfactory report card grades in health habits, work habits, and language as well as unsatisfactory scores on reading readiness status and reading grade placement were found more frequently within the speech handicapped sample.

4. Aggression, fears and anxieties, and perception of parental figures as authoritarian appeared significantly more often in the CAT protocols of the articulation sample.

5. The mean number of positive CAT protocol outcomes for the normal sample was significantly greater than that for the articulation group.

6. The average total oral participation scores of the normal group were significantly greater than those of the speech handicapped group.

The present investigation suggests the existence of a relationship between a diffuse, nonorganic articulatory disorder and psychosocial factors, indicating a need for additional research. These findings preclude the efficacy of an essentially symptomatic treatment approach in speech therapy with children who present diffuse, nonorganic articulatory problems.

References

1. Bellak, L., & Bellak, Sonya S. *Children's Apperception Test.* New York: C.P.S., 1946.
2. Blau, A., & Hulse, W. C. Preventive psychiatry in the care of the newborn and infant. Paper read at Amer. Psychiat. Ass., Los Angeles, May, 1953.
3. Doll, E. A. *Vineland Social Maturity Scale.* Philadelphia: Educational Test Bureau, 1948.
4. Hildreth, Gertrude H. *Metropolitan Achievement Tests.* New York: World Book, 1948.
5. Hildreth, Gertrude H. *Metropolitan Readiness Tests.* New York: World Book, 1949.
6. Kuhlmann, F., & Anderson, Rose G. *Kuhlmann-Anderson Test.* (6th Ed.) New Jersey: Personnel Press, Inc., 1952.
7. McCarthy, Dorothea. Language disorders and parent-child relationships. *J. Speech Hearing Disorders,* 1954, 19, 514-523.
8. Mowrer, O. H. Speech development in the young child: 1. The autism theory of speech development and some clinical applications. *J. Speech Hearing Disorders,* 1952, 17, 263-268.
9. Powers, Margaret H. Functional disorders of articulation—symptomatology and etiology. In L. E. Travis (Ed.), *Handbook of speech pathology.* New York: Appleton-Century-Crofts, 1957. Pp. 707-768.

10. Pronovost, W. L. A survey of services for the speech and hearing handicapped in New England. *J. Speech Hearing Disorders,* 1951, 16, 148-156.

11. Read, Elizabeth C. A teacher administered rating chart for talks at the primary level. Unpublished Master's thesis, Boston Univer., 1953.

12. Silverberg, W. V. *Childhood experiences and personal destiny.* New York: Springer, 1952.

13. Spriesterbach, D C. Research in articulation disorders and personality. *J. Speech Hearing Disorders,* 1956, 21, 329-335.

14. Travis, L. E. A point of view on speech correction. *Quart. J. Speech,* 1936, 22, 57-61.

SELF-CONCEPTIONS, MOTIVATIONS, AND INTERPERSONAL ATTITUDES OF EARLY- AND LATE-MATURING GIRLS[1]

MARY COVER JONES *and* PAUL HENRY MUSSEN
University of California, Berkeley

"The changing body and the changing self" is a phrase associated with adolescent development (18). It suggests that the shaping into mature form of the childhood body pattern is accompanied by new self-concepts. These altered attitudes toward the self reflect at least in part the youth's response to his physical metamorphosis.

What "growing-up" connotes for the individual adolescent depends upon a complex of psychobiological factors. One of the most important of these is rate of physical maturation. Adolescent growth may be relatively regular and even, or it may be uneven or abrupt. The timing of puberty, in relation to social norms of the peer group, may present problems of special importance for some adolescents.

Previous reports of systematic comparisons between the behavior and personality characteristics of early- and late-maturing adolescents have indicated that acceleration in growth tends to carry social advantages for boys (9, 10) but disadvantages for girls (7). At their peak of growth, early maturing girls are not only taller than their girl classmates but are actually taller than most of the boys in their class (14). They are conspicuously large at a time when physical size is not an asset for girls in our culture. Many girls consider tallness to be a physical stigma (15). At the end of adoles-

[1] The TAT data for this study were obtained by Harold E. Jones in connection with a test program at the Institute of Child Welfare.

CHILD DEVELOPMENT, Vol. 29, No. 4 (December, 1958)

cence the early-maturing are no longer taller than their age-mates, but in body proportion they tend to have a broad and stocky build (13), less attractive (in terms of current feminine standards) than the more slender physique of the late-maturing.

Among boys, ascendance in size and musculature is an asset because of our cultural values and the functional advantages of such a build for athletic prowess. This more favorable status is indicated in observational records for early-maturing boys. Staff members rated them as physically more attractive and better-groomed than the late-maturing, and in social situations they were more poised and matter-of-fact, and less attention-seeking (9).

In contrast, both classmates (16) and adult observers (12) saw the early-maturing girls as relatively submissive, listless or indifferent in social situations, and lacking in poise (7). Such girls have little influence upon the group and seldom attain a high degree of popularity, prestige or leadership (8).

The girls in the slower-maturing classification were seen as relatively more outgoing and more assured. They were eager, animated, peppy, and talkative. This behavior seems to be acceptable among girls since those who exhibit it are also described as confident and having leadership abilities (7). While the same characteristics of expressiveness are attributed to slow-growing boys, it is associated in their case more specifically (and especially in later adolescence) with show-off behavior, affectation, and tenseness (6, 9).

In accounting for these sex differences in the response to early or late puberty, we may note that although early-maturing boys have physical advantages over other boys and are socially in step with girls, the girl who develops earlier than her classmates may be temporarily isolated. H. E. Jones has expressed this as follows:

> The early-maturing girl quite naturally has interests in boys and in social usages and activities more mature than those of her chronological age group. But the males of her own age are unreceptive, for while she is physiologically a year or two out of step with the girls in her class, she is three or four years out of step with the boys—a vast and terrifying degree of developmental distance (7, p. 78).

A study of responses to the Thematic Apperception Test, given to members of the Adolescent Growth Study when they were seniors in high school, yielded a somewhat unfavorable psychological picture for the late-maturing boys. Compared with their early-maturing peers, they showed greater evidence of negative self-concepts, prolonged dependency needs, feelings of rejection by others, rebellious attitudes toward parents, and strong affiliative needs. These findings were in agreement with evidence from other sources (10).

A similar TAT comparison of early- and late-maturing girls should be expected to show results different from those obtained for boys. Thus, it

might be expected that early-maturing girls would reveal negative self-feeling and less satisfactory interpersonal attitudes.

Procedure

The present study, paralleling that for boys, was designed to investigate the relationship between maturational status and self-conceptions, motivations, and interpersonal attitudes in a normal public school sample of girls (5). Personality assessment was made on the basis of their responses to the Thematic Apperception Test (TAT).

The 34 17-year-old girls of this investigation constitute approximately the 20 per cent at each extreme of the total sample of the Adolescent Growth Study, selected on the basis of their physical maturity status as determined by X-rays of the wrists and hands (1). Sixteen had been among the most consistently accelerated over a four-year period during adolescence; the other 18 were among the most consistently retarded. All of the subjects took the TAT at around age 17 when they were seniors in high school.

The TAT consisted of 18 pictures: nine from the Murray set which is now standard (cards 1, 5, 6, 7BM, 10, 11, 14, 15, 17); five pictures from the set generally used in 1938 when these data were collected (a man and woman seated on a park bench; a bearded old man writing in an open book; a thin, sullen, young man standing behind a well-dressed older man; a tea table and two chairs; an abstract drawing of two bearded men); and four cards not in the Murray series (a madonna and child, the nave of a large church, a dramatic view of mountains, a boy gazing at a cross which is wreathed in clouds).

The tests were administered individually. Each card was projected on a screen while the subject told a story which was recorded verbatim. Standard instructions were given for the Murray cards, and subjects were asked to describe the feelings elicited by the other four pictures. Most of the stories were brief.

The scoring scheme involved counting the relevant needs, press, and descriptions of the heroes of the stories, the assumption being that the storyteller has identified with the hero; the hero's needs are the same as the girl's; the press that impinge upon the hero are the ones that affect the girl telling the story. A total of 20 needs, press, and descriptive categories, each defined as specifically as possible, was developed in the analysis of the protocols. A score for each subject for each TAT category was derived by counting the number of stories in which it appeared. Table 1 presents a list of the categories used, together with definitions of these categories.

To test the reliability of this analysis, one of the authors (PM) and another psychologist[2] independently scored 15 complete protocols (300

[2] We are indebted to Dr. Virginia B. Ware for her participation in this aspect of the study.

stories). The percentage of interrater agreement was 90, computed by the usual formula (number of agreements divided by number of agreements plus number of disagreements).

In order to eliminate bias, the scoring used in the present study was done "blind," that is, independently of knowledge of the subject's maturational status.

Results

Frequency distributions of the scores of all subjects were made for all the TAT variables. Each distribution was then dichotomized at the point which most nearly enabled the placing of half of the 34 subjects above, and half of them below, the dividing point. Subjects having scores above this point were considered high in this particular variable; those with scores below this point were considered low in this variable.

Table 1 lists the TAT variables together with the number of late- and early-maturers with high scores in each variable. The exact probabilities of obtaining these distributions of high and low scores in the two groups (or all other possible more extreme sets), calculated in accordance with Fisher's method (3), are given in the fourth column. The last column gives the levels of significance of the differences between early- and late-maturing boys on these same variables, previously reported by the authors (10).

As may be seen from Table 1, early- and late-maturing boys differed from each other on many more characteristics than the two groups of girls did. The boys' groups were significantly different from each other, at the 5 per cent level or better, in six of the 20 variables scored, while the early- and late-maturing girls differ significantly in only two of the variables (*negative characteristics* and *n Recognition*). It should be noted, however, that the *direction* of the differences tended to be the same, rather than reversed, in the two sets of data. For example, the following similarities may be noted:

1. In this list of characteristics a significantly greater proportion of late-maturing girls than of early-maturing girls have high scores on *negative characteristics*. This finding is similar to that found in the comparison of early- and late-maturing boys. For girls, it is contrary to expectation.

2. The differences between early- and late-maturing girls in respect to *p Dominance* and *p Rejection* are similar to those for early- and late-maturing boys in these variables. These may be interpreted to indicate slightly poorer parent-child relationships among the late-maturing.

3. The early- and late-maturing boys differ significantly on *n Autonomy 1*, suggesting a greater tendency for the late-maturing to avoid or defy authority. The differences are in the same direction for girls, but are not significant.

4. Similar results for boys and girls in *n Succorance* may be interpreted as showing some tendency for stronger dependency needs in the late-maturing.

TABLE 1

NUMBER OF EARLY- AND LATE-MATURERS SCORING HIGH IN TAT VARIABLES

TAT Variable (with Definition)	*High Early-Maturers*	*High Late-Maturers*	*p*	*p* (boys)
Negative Characteristics—H is described in negative terms (e.g., imbecile, weakling, fanatic)	3	13	.002	.01
p Dominance 1—H forced by parents to do something he doesn't want to	2	7	.08	.09
p Dominance 2—H prevented by parents from doing something he wants to	3	6	—	—
p Dominance 3—Total instances of H's being forced by parents to do something and/or prevented from doing something	4	10	.07	.11
p Rejection—H rejected, scorned, or disapproved by parents or authorities	3	7	.21	.03
n Aggression 1—H is aggressive in physical, asocial way	8	9	—	.02*
n Aggression 2—H is mad at someone, argues	12	14	—	.10*
n Aggression 3—Total of all H's aggressive actions	5	8	—	.10*
n Autonomy 1—H leaves home	8	10	—	.20
n Autonomy 2—H disobeys or defies parents	4	5	—	.11
n Autonomy 3—Total of instances in which hero leaves and/or defies his parents	5	7	—	.02
n Affiliation 1—H establishes good relations with his parents	9	8	—	—
n Affiliation 2—H falls in love, has a romance, marries	11	7	.08	.05*
n Affiliation 3—Total instances in which H establishes and/or maintains friendly relations	7	6	—	.11*
n Succorance—H feels helpless, seeks aid or sympathy	5	10	.16	.06
p Nurturance 1—H is helped, encouraged, or given something by parents	5	7	—	.18
p Nurturance 2—H is helped, encouraged, or given something by someone else (not parents)	5	10	.16	.02
n Achievement—H attempts to attain a high goal or do something creditable	7	13	.09	—
n Recognition—H seeks fame and/or high prestige status	3	12	.01	—
Denial of Feeling—*S* states that picture elicits no thoughts or feelings	10	9	—	.06

* Differences are in the opposite direction for boys as compared with girls.

5. Similar results for boys and girls in *p Nurturance* (significant in one variable for boys) may also be interpreted as indirect indications of stronger dependency needs in the late-maturing.

The chief differences between the sexes are as follows:

1. With respect to *n Aggression* more early- than late-maturing girls show "argumentative aggression," but the two groups of girls do not differ in physical aggression. On the other hand, more early-maturing than late-maturing boys show high degrees of both kinds of aggression.

2. On one category of *n Affiliation* (involving romantic love) higher proportions of high scores are shown for early-maturing girls as contrasted with their late-maturing peers. The differences between the early-maturing and the late-maturing boys are in the opposite direction for this category.

3. The variables *n Achievement* and *n Recognition* do not differentiate the two groups of boys. Among girls scores are higher for the late-maturing, very significantly so in the case of *n Recognition*.

4. *Denial of feeling* does not differentiate early- and late-maturing girls but tends to yield higher scores for early-maturing boys.

Discussion

The failure of the TAT data to support observational findings, especially with reference to the variable, negative characteristics, might be accounted for in a number of ways. Some writers report that in many cases thematic fantasies and manifest behavior operate independently and are even negatively related (11). If we assume this to be the case for our subjects, no further explanation would be needed. But there is also evidence from the literature that, for some groups, TAT findings and overt behavior may be congruent (4). Our data on boys are in line with this assumption, since, according to observational ratings, late-maturing boys tend to be socially disadvantaged, and, according to the TAT, personally more maladjusted.

The findings for girls are quite different, however. The early-maturing received more unfavorable ratings from both peers and adult observers on many characteristics. But in the TAT they appear to be somewhat better adjusted than their late-maturing peers. This discrepancy between observers' ratings and the picture derived from the personality tests may stem partly from the fact that the reported observational records represented an average of repeated ratings taken over a period of time (from 11 to 17 years) while the TAT stories were collected at the end of this period.

Girls who enter puberty early would be expected to have more difficulties in personal-social relations when they are out of phase with their group. However, after the peer group "catches up," these difficulties would be reduced. By the end of senior high school maturational discrepancies, and social distance due to this factor, would be less marked. It is also possible that even a slight improvement in status would bolster morale and be

reflected in a projective technique designed to register attitudes and self-concepts.

There is some slight evidence of a trend toward improved social status for the early-maturing in observational ratings over the seven-year period. Twenty-five items concerned with appearance, emotional tone, social participation, responsiveness, and assurance were used in the comparison. Three of these reflected an improved status at the twelfth grade level for early-maturing girls (2). In two of these, "laughing" vs. "sober" and "sociable" vs. "unsociable,'" the accelerated girls, while still rated lower than the late-maturing, had improved sufficiently so that the differences between the two groups were no longer significant. But for one important characteristic, "popular" vs. "unpopular," the average ratings for the accelerated girls were now actually slightly higher than for the late-maturing, though the differences were not significant. This last year at high school was the only period when the early-maturing girls were rated by observers as above average in popularity.

It is conceivable that other improvements in social relationships were undetected because of the "halo effect" which, in spite of precautions, may have influenced observers who had rated these same adolescents in earlier years. It is not unlikely that if these girls of more mature status had been observed in social groups of their own choosing (presumably outside of school) the behavior picture might have been more favorable.

It may be noted that over the seven-year period the observational records received little corroboration from a self-report inventory (17). Although differences were not consistent in all categories, the early-maturing girls tended to score more favorably than the slow-maturing on "total adjustment," and also on family adjustment and feelings of personal adequacy. These data from the self-report inventory seem to be generally consistent with the findings from the TAT.

However, we may note that in both the inventory and the TAT the early-maturing girls appear in a somewhat better light than in their reputation scores or in ratings by adult observers. In some individual cases a favorable self-report score should not be taken at face value, in view of the tendency for some individuals to cover up or deny their deficiencies.

The only other variable which yields a significant difference between the maturity groups is the category *n Recognition,* late-maturing girls manifesting a greater desire for personal recognition. The results for *n Achievement,* though not showing significant differences, tend to support these findings. Other data for this group of girls would lead us to expect this relationship between maturity status and desire for recognition. Late-maturing girls were rated by adult observers as attaining higher prestige, showing more leadership, and having greater stimulus value than their early-maturing peers (7). They were also mentioned more frequently in the high school daily paper over a three-year period and were elected to more offices in extra-curricular activities (8). The late-maturing girls' lead-

ership abilities, their greater social participation, and their apparent social success may have been more closely related to desires for recognition and achievement in the social sphere than to a need for affiliation (8).

It should be noted that, among boys, *n Achievement* and *n Recognition* were not significantly associated with rate of physical maturation. Perhaps this is due to the fact that for boys in our culture the pressures to strive for achievement and personal recognition are powerful and pervasive; hence, the boy's physical status may have little influence on his acquisition of strong achievement and recognition needs. Since these cultural pressures are undoubtedly less severe for girls, the strength of these personal needs may be more influenced by such factors as rate of physical maturation.

As we have pointed out in an earlier article (10), the relationship between physical status and psychological characteristics in boys is by no means simple. The evidence of the present study indicates that this relationship is even more complex in the case of girls. While the TAT analysis reported in this study suggests that early-maturing girls have fewer negative self-concepts and fewer needs for personal recognition, the results must be interpreted very cautiously. Since only two variables were found to be significantly related to physical status, it is obvious that many psychological and social factors are more important than rate of maturing in determining girls' self-concepts and personality characteristics. Furthermore, these data, considered together with the data from earlier studies on girls (7, 8), suggest that the rate of maturation may affect overt behavior and covert characteristics in different—sometimes seemingly contradictory—ways.

It is also possible that, at least for girls, early- or late-maturing means different things at different stages of adolescent development. It has been proposed that since girls who enter puberty early are out of step physically with both the boys and girls in their classrooms, they tend to be socially handicapped during early adolescence. We have assumed that this would carry emotional hazards, and evidence is available from observational data and reputation measures to indicate that this is the case (7).

However, the accelerated girl may gain assurance from knowing that she is on the way toward a goal which is a common task for all adolescents, that of being an adult. By the end of high school, many girls in this group were beginning to feel that they had made satisfactory progress toward this goal. If, in addition to this, she can cope with the problems of this period without too much stress, her self-esteem and feelings of adequacy may be enhanced. A resulting improvement in self-concepts may be reflected in the relative infrequency of negative characteristics in TAT stories.

In conclusion, it is evident that each individual's unique personality structure is determined by a complex of interacting variables, including rate of maturation. Comments made by these subjects as young adults indicate that they were aware of a variety of surface phenomena which affected their adolescent adjustment:

"High school is not a pleasant memory. I felt remote from my mother. If I could have talked to her, it would have helped" (a slow-maturer).

"I wasn't very happy in adolescence. My father was out of work. I felt inferior outside my own circle of friends—I always aimed to please" (a very popular late-maturing girl).

"I was slightly rattle-brained" (a popular late-maturing girl).

"I didn't have much fun in high school. I look forward to more happiness now than I did when I was in high school. I was an ugly duckling" (a slow-maturer who ascribed many negative characteristics to the hero).

"I seemed to be separated from friends in high school. I'm more outgoing now, less cautious and fearful" (accelerated girl).

"I was overweight and sensitive about it—now I take things more for granted" (accelerated girl).

"I had a feeling of being different when growing up" (accelerated).

"I felt stupid in school" (accelerated girl).

"I was very lacking in self-confidence in high school" (accelerated girl).

"I'm more optimistic now. I didn't know many people in high school. I would make an effort to get on with people if I had it to do over again" (accelerated).

Feelings of inadequacy and isolation are expressed by these girls and they are attributed to lack of mental ability, financial difficulties, separation from parents, poor social status, overweight, and unattractiveness. They are about equally common among those whose maturational status was at one extreme or the other.

It is obvious that the findings for this specific group of girls need to be particularized for each individual. These results might be modified also for girls in another geographical area or social level or in another generation. It is possible that school and community programs may be able to de-emphasize maturational status by providing an easier access to mixed social groups through classroom, extra-curricular, and recreational activities which cut across age classifications.

Summary

The present study was designed to investigate the relationship between maturational status and TAT scores for a group of physically-accelerated as contrasted with a group of slow-developing girls from a normal classroom sample. The TAT protocols of 34 17-year-old girls—16 who had been consistently accelerated and 18 who had been consistently retarded—were analyzed according to a scoring scheme involving 20 needs, press, and descriptive categories.

The scores of early- and late-maturing in each of the categories were compared. Earlier reports (7, 8) had indicated that girls who reach puberty early are likely to be socially disadvantaged, at least until the rest of their age group "catch up" with them. It was assumed that this social disadvantage would be reflected in the TAT protocols and that differences

between the two maturity groups in self-concepts, attitudes, and motivations would be found. Analysis of the data of the present study found few striking differences between the two groups of girls. However, early-maturing girls had significantly lower scores on the category *negative characteristics,* indicating more favorable self-concepts. This finding is contrary to what might have been expected on the basis of observational ratings by adults and reputational ratings by classmates. On the other hand, the TAT results are in line with scores (total adjustment, self-adequacy, family adjustment) on a self-report inventory.

Late-maturing girls have significantly higher scores on *n Recognition,* which is corroborated by data from other sources.

When the differences between early- and late-maturing girls are compared with the differences between early- and late-maturing boys (10), they are found to be in the same direction more often than in the opposite. These findings are interpreted to indicate that late-maturing adolescents of both sexes are characterized by less adequate self-concepts, slightly poorer parent-child relationships, and some tendency for stronger dependency needs.

It has been emphasized that complex psychological and cultural factors as well as maturational status contribute to personality development and that the pattern of these influences varies for each individual.

References

1. Bayley, N. Size and body build of adolescents in relation to rate of skeletal maturing. *Child Develpm.,* 1943, 14, 51-89.
2. Everett, E. G. Behavioral characteristics of early- and late-maturing girls. Unpublished Master's thesis, Univer. of California, 1943.
3. Fisher, R. A. *Statistical methods for research workers.* (7th Ed.) Edinburgh: Oliver & Boyd, 1938.
4. Harrison, R. The Thematic Apperception Test. In L. K. Frank *et al.,* Personality development in adolescent girls. *Mongr. Soc. Res. Child Develpm.,* 1951, 16, 60-88.
5. Jones, H. E. Observational methods in the study of individual development. *J. consult. Psychol.,* 1940, 4, 234-238.
6. Jones, H. E. *Development in adolescence.* New York: Appleton-Century, 1943.
7. Jones, H. E. Adolescence in our society. In *The family in a democratic society: anniversary papers of The Community Service Society of New York.* New York: Columbia Univer. Press, 1949, Pp. 70-82.
8. Jones, M. C. A study of socialization at the high school level. In press.

9. JONES, M. C., & BAYLEY, N. Physical maturing among boys as related to behavior. *J. educ. Psychol.*, 1950, 41, 129-148.

10. MUSSEN, P. H., & JONES, M. C. Self-conceptions, motivations, and interpersonal attitudes of late- and early-maturing boys. *Child Develpm.*, 1957, 28, 243-256.

11. MUSSEN, P. H., & NAYLOR, H. K. The relationship between overt and fantasy aggression. *J. abnorm. soc. Psychol.*, 1954, 49, 235-240.

12. NEWMAN, F. B. The adolescent in social groups. *Appl. Psychol. Monogr.*, 1946, 9.

13. SHUTTLEWORTH, F. K. Sexual maturation and the physical growth of girls age 6 to 19. *Monogr. Soc. Res. Child Develpm.*, 1937, 2.

14. SHUTTLEWORTH, F. K. The physical and mental growth of girls and boys age 6 to 19 in relation to age at maximum growth. *Monogr. Soc. Res. Child Develpm.*, 1939, 4.

15. STOLZ, H. R., & STOLZ, L. M. Adolescent problems related to somatic variations. *Yearb. nat. Soc. Stud. Educ.*, 1944, 43, Part 1. Pp. 80-99.

16. TRYON, C. M. Evaluations of adolescent personality by adolescents. *Monogr. Soc. Res. Child Develpm.*, 1939, 4.

17. TRYON, C. M. *U. C. Adjustment Inventory I: Social and emotional adjustment.* Berkeley: Univer. of California Press, 1939.

18. ZACHRY, C. B. *Emotion and conduct in adolescence.* New York: Appleton-Century, 1940.

A REINTERPRETATION OF THE BARKER, DEMBO, AND LEWIN STUDY OF FRUSTRATION AND REGRESSION

JOHN M. DAVIS
Yale University School of Medicine

Barker, Dembo, and Lewin's (BDL) classic study of frustration and regression (1) has been widely discussed in psychological literature. Child and Waterhouse (CW) reinterpreted the experiment and have come to conclusions which differ from those of the original authors (2). In the present paper further analysis of the data will be presented in an attempt to evaluate the two interpretations.

BDL feel that psychological development can be characterized by structural changes in the life space such that there is increased differentiation of the person into subsystems and a change of organization of the person so that the subsystems are hierarchically related to one another rather than related by simple interdependence. BDL define regression as a primitivization of behavior: the subject's behavior becomes less organized and more like that of a less mature person. The BDL definition of regression differs therefore from the psychoanalytic definition in that the behavior elicited after frustration need not be specific responses that have been earlier manifested and abandoned by the organism. In the experiment 30 children (ages 3 to 5) were allowed to play with "average" desirable toys during a free play period. The next day they were allowed to play with some *very desirable* toys, but then both the toys of the first period and the children were moved to another part of the room and a wire screen partition was lowered so the children could see the very desirable toys, but could play *only* with those toys they had played with in the first situation. During both the free play and the frustration period the subject's behavior was recorded; later it was analyzed into a number of categories. The subject's play was rated on a constructiveness scale. The constructiveness scale measured the degree of organization, maturity, and differentiation of the play. It correlates highly with mental age and chronological age. The constructiveness of the play of children in the frustration situation was considerably lower than that in the free play situation.

In the analysis of the behavior of the children BDL differentiate between primary and secondary play. In primary play the child seems (to the observer) to be entirely absorbed in the play; in secondary play the child is playing but also doing something else at the same time. The constructiveness of secondary play is considerably lower than that of primary play.

Barker, Dembo, and Lewin interpret the experimental results in terms of their theoretical position outlined above. They feel that tension caused

CHILD DEVELOPMENT, Vol. 29, No. 4 (December, 1958)

by the "frustration" results in a dedifferentiation of the life space. According to Lewin's theory the degree of differentiation of a whole is inversely proportional to the strength of the pressure of the tension, when the tension is greater than limits set by the strength of the boundary of the natural parts of the whole. Lewin also feels that a decrease in the variety of behavior must occur when a part of the whole is kept in a fixed state because of the frustration.

Child and Waterhouse (CW) give the experiment a different interpretation (2). They point out that BDL's interpretation of the data in terms of dedifferentiation is circular, since no criterion for differentiation is present. They point out further that there is not, as BDL claim, a decrease in the variety of behavior during the frustration period; there is rather, in direct contradiction to BDL, an increase in the variety of behavior during the frustration period.

CW interpret the BDL study in terms of interference. That is, the lowered constructiveness of play in the frustration situation is explained by assuming that there is something about the frustrating situation which causes the subjects to make responses which interfere with the play responses. To illustrate the interference hypothesis let us take an example from another situation: Let us assume that there is a decrease in the quality of a man's driving when he is both driving the car and smoking a cigar, instead of driving the car only; that is, the responses of smoking the cigar interfere with those of driving the car.

CW presents the following evidence for the interference hypothesis: (a) During the frustration period, there was three times as much time devoted to nonplay activities as to play activities, and there was a greater proportion of secondary play. (b) With chronological and mental age partialed out, there is a correlation of -0.72 between the combined categories of escape and barrier behavior during the frustration period and constructiveness of play during this period. (c) With chronological and mental age partialed out, there is a correlation of -0.83 between constructiveness in the free play period and the combined categories of activity with the experimenter, island behavior, window behavior, and looking and wandering around during the free play period. (d) There is no significant difference in the constructiveness of secondary play between the play and the frustration period. This result would be predicted by the CW analysis because, in secondary play, there is already interference, so the extra interference in the frustration period would have no effect. The BDL interpretation cannot explain this fact.

However, it would be possible for BDL to counter the CW interpretation by some of the following counterarguments. BDL could argue in relation to CW points (a) and (b) that frustration "causes" the nonplay behavior and, therefore, the more frustrated subjects would engage in more nonplay behavior, i.e., more escape and barrier behavior. On point (c) BDL could argue that their theory makes no prediction in regard to indi-

vidual differences in the free play situation. Hence, the CW interpretation seems more reasonable intuitively, but their argument could be countered by the BDL group. What is needed to differentiate the two interpretations is more evidence.

Problem

In their original paper, BDL give the score each subject made in each content category (1, Table 1, p. 85; Table 6, p. 119; Table 7, p. 123; Table 8, p. 133; p. 135; Table 9, p. 142; Table 20, p. 191; Table 21, pp. 191-192). It is possible, therefore, to go to these original data and analyze them in new ways.

The BDL interpretation and the CW interpretation would predict different correlations between these content categories. According to the BDL interpretation, those categories which are related to the strength of frustration would be negatively related to constructiveness. The BDL interpretation would therefore predict that the categories of mood satisfaction–unsatisfaction index[1] and escape behavior would correlate negatively with constructiveness. The CW interpretation would predict that those categories which imply the strongest interference states would correlate negatively with constructiveness, specifically, that the category of barrier behavior would correlate negatively with constructiveness. Barrier behavior would imply a strong interference state because it indicates that the subjects were concerned with the very desirable toys behind the barrier.

Method

Rank order correlations were calculated from the scores of each subject on each of five content categories. Each category was correlated with every other category. The matrix of correlations is listed in Table 1. The reliability of the various behavioral rating categories could not be calculated from the data available in the BDL report.

Results

An inspection of the table of correlations shows that only those correlations predicted by the CW interpretation correlate significantly with constructiveness. Barrier behavior correlates −.39 with constructiveness and −.50 with maximum constructiveness. None of the correlations predicted by the BDL interpretations was significant: escape behavior correlated with constructiveness (−.17) and maximum constructiveness (−.06), and mood index correlated with constructiveness (+.04) and maximum con-

[1] BDL construct the mood satisfaction–unsatisfaction index from behavioral ratings of the following categories: singing and humming, friendly conversations, happy actions, unhappy actions, restless actions, knocking and kicking. The criteria for these categories and the construction of the index are discussed by BDL (1, pp. 187-204).

structiveness (+.29). In short, the evidence from the correlational analysis supports the CW interpretation and does not support the BDL interpretation.

TABLE 1

INTERCORRELATIONS OF BEHAVIOR RATINGS

	C	*MC*	*B*	*E*	*MI*
C		.48*	—.39*	—.17	.29
MC	..		—.50*	—.06	.04
B	..	..		—.08	—.23
E	..	..	..		—.24

NOTE.—C, constructiveness; MC, maximum constructiveness; B, barrier behavior; E, escape behavior; MI, mood index.

* Significant at .05 level.

SUMMARY

Child and Waterhouse have reinterpreted the much descussed frustration and regression experiment of Barker, Dembo, and Lewin and have arrived at different conclusions from those of the original authors. Barker, Dembo, and Lewin rated the behavior of children in a free play situation and in a frustration situation on a number of content categories. They found that the play of the children was less constructive during the frustration period. They interpreted the results in terms of the Lewinian primitivization. CW interpreted the experiment in terms of an interference hypothesis. In an attempt to differentiate the two interpretations, five content categories were correlated with each other, using data in BDL's original paper. From the interpretations of BDL and CW, different predictions were deduced about the correlations between categories. The correlations confirm the prediction from the CW interpretation and do not confirm that from the BDL interpretation. Therefore, the CW interpretation seems to be more consistent with the data.

REFERENCES

1. BARKER, R. F., DEMBO, T., & LEWIN, K. Frustration and regression: an experiment with young children. *Univer. Ia. Stud. Child Welf.*, 1941, 18, 1-314.
2. CHILD, I. L., & WATERHOUSE, I. K. Frustration and the quality of performance. 1. A critique of the Barker, Dembo, and Lewin experiment. *Psychol. Rev.*, 1952, 59, 351-362.

DOMINANCE ATTITUDES IN PARENTS AND ADULT AVOIDANCE BEHAVIOR IN YOUNG CHILDREN[1]

E. PHILIP TRAPP *and* DONALD H. KAUSLER
University of Arkansas

The fruitfulness of predicting child behavior from family variables is indicated by the trend of research in this direction (1, 4). In most of these studies, however, the effort has been to relate children's behavior to parental practices rather than to parental attitudes, mainly because of the general unreliability of attitude measurements in this area. The advantage of being able to predict children's behavior from measures of parental attitudes is obvious in terms of convenience and simplicity of data gathering.

Read (12) attempted to relate parents' expressed attitudes to nursery school children's ratings on a behavior inventory and found some relationship between attitudes of parental control and ratings on the inventory. Children of parents favoring strong control were rated with more undesirable behavioral characteristics than children of parents expressing a more liberal attitude. The ratings were the pooled judgments of the nursery school teachers.

The present study is a further investigation into the effects of parental attitudes of dominance on child behavior. Shoben (15) has developed a parent attitude scale which shows promise of being a reliable and sensitive instrument, although a recent validity study (7) indicates a limitation in its

[1] Special thanks are reserved for the nursery school staff at the University of Arkansas for their kind assistance throughout the study, and for Miss Peggy Wall, who made the observations on the children.

clinical use. The three major attitudes it purports to measure fall under the subscale headings of Dominant, Possessive, and Ignoring. Thus far, the sensitivity of the scale has been demonstrated in differentiating mothers of problem children from mothers of nonproblem children. The usefulness of the scale in differentiating less extreme cases, that is, parents of children not undergoing psychiatric treatment, is tested in the present study. More specifically, the focus of this study is on the relationship between parental dominance attitudes and adult avoidance behavior in young children.

Children of parents reflecting strong dominance attitudes are likely to develop more than the average amount of negativistic feelings toward their parents. These feelings should reinforce parent avoidance responses and generalize to other adult relationships. It is also logical to propose that parents with weak dominance attitudes might increase negativistic feelings in their children. In such cases, the child, because of insufficient guidance, is likely to experience more than the ordinary amount of frustrations and project his frustration on his parents. One manifestation of this process would be in increased parent avoidance behavior which would generalize to other adult relationships.

This proposed curvilinear relationship between parents' dominance attitudes and children's adult avoidance behavior is predicted to be in evidence only with children displaying normal personality development. In the more extreme cases of problem children, the higher anxiety drives would probably generate more variable response patterns, reflected in either excessively dependent behaviors or extreme forms of withdrawal behaviors. The net result would obscure the specific effects hypothesized here.

As a second and related problem, the present study investigated the relationship between differences in expressed parental dominance attitudes and children's behavior. It is proposed that parents with large differences in dominance attitudes between themselves would create a conflict situation for the child. Therefore, the parent-child relationships within these families are likely to be less positive than parent-child relationships within families where there is more agreement on dominance attitudes. Evidence of this less positive relationship should be indicated by increased adult avoidance behavior in their children.

These two predictions regarding the relationship between parental dominance attitudes and children's adult avoidance behavior may be stated more precisely in the following two hypotheses: (a) Nursery school children of parents scoring either relatively high or low on a dominance attitude scale will avoid adult contacts in a free-activity setting to a greater degree than nursery school children of parents scoring in the intermediate range on a dominance attitude scale. (b) Nursery school children of parents reflecting large differences in scores on a dominance attitude scale will avoid contacts in a free-activity setting to a greater degree than nursery school children of parents reflecting small differences in scores on a dominance attitude scale.

Method

Subjects

At the outset of the study, 20 children attending the University of Arkansas Nursery School served as *Ss*. By the end of the study, the normal process of attrition left completed data on 16 children. Eight *Ss* were males, ranging in age from 3-4 to 5-0, with a mean age of 4-2 (computed in years and months). The eight female *Ss* also had an age range from 3-4 to 5-0, with a mean age of 4-4. The educational and socioeconomic backgrounds of the parents were fairly homogeneous, most of them of professional middle class.

Procedure

Each parent was given the USC Parent Attitude Survey (13). The Dominant subscale served as the predictor measure. The parents were informed that the survey was part of a nation-wide study to investigate sectional differences in child-training attitudes. They seemed satisfied with the explanation and cooperated fully with the project. Anonymity was assured them by the use of coded numbers in place of their signatures. The survey was administered and collected by the director of the nursery school, who had known each parent for at least six months.

Each child was observed individually at five-minute time intervals, approximately once every 10 days through a period of 15 weeks, resulting in a total of 50 minutes of observation per child. The children were observed in the mornings between 11:00 and 11:50, the free-activity period in the nursery school program. The observations followed a prearranged schedule to insure that each child was observed throughout the entire period and never more than once at the same time of the period.

The observer was equipped with two concealed stop watches. One ran continuously for the five-minute time interval and the other watch ran only at the time the child was in contact with an adult. At the end of a day's observation, the observer made notations on the nature of the adult contacts and other qualitative remarks on the children's behavior. Since the children were accustomed to adults about them (the nursery school provides training facilities for University students), they did not show any unusual curiosity toward the observer, and, in fact, gave no indication that they were aware of being under any special form of observation. To assist in achieving this desired end, the observer spent a few days in the play area before beginning the formal observations.

Results

Regarding the first hypothesis investigated, the two dominance scores for each set of parents were added together and a combined single dominance score determined. These combined scores ranged from 233 to 304.

The parents were divided into quartiles based on their combined scores. The first and fourth quartiles represent extreme dominance attitudes; the second and third quartiles represent intermediate dominance attitudes.

Children with parents in the extreme quartiles constituted one group of *Ss* (Extreme Group, $N = 8$); children with parents in the intermediate quartiles constituted a second group (Intermediate Group, $N = 8$). The two groups were found to be nearly identical in average age and average length of attendance in the nursery school. To test the related hypothesis, the mean time (in minutes) spent in adult contact was determined for each group, and the difference between group means was tested for significance by the *t* test. Summary data for this analysis appear in Table 1.

TABLE 1

SUMMARY OF CHILDREN'S AVOIDANCE BEHAVIOR: GROUPS DETERMINED BY COMBINED PARENTAL DOMINANCE SCORES

Group	*Mean* (min.)	*SD*	*t*
I Children of parents with extreme attitudes ($N=8$)	5.7	4.6	2.3*
II Children of parents with intermediate attitudes ($N=8$)	12.5	6.3	

* $p < .05$ (two-tailed test).

The results support the hypothesis ($t = 2.3$, $p < .05$, two-tailed test). Children with parents scoring either high or low on a dominance attitude scale displayed more adult avoidance behavior than children with parents in the intermediate range on the scale. A clearer pattern of this curvilinear relationship may be seen in a further breakdown of the data in Table 1 in terms of the quartile means. The means of the quartiles from top to bottom were 6.5, 8.6, 16.4, and 4.9, respectively.

Regarding the second hypothesis investigated, a difference score on dominance attitude for each set of parents was obtained in the following manner. For each of the 34 items on the USC Parent Attitude Survey measuring dominance, a score was recorded indicating the discrepancy between the two parents on that item. The 34 discrepancy scores were added to produce the total score for each set of parents. The parents were again divided into quartiles, this time on the basis of their difference score. The difference scores ranged from 18 to 33.

Children with parents in the first quartile (large difference between parents on dominance attitude) constituted one group of *Ss*; children with parents in the fourth quartile (small difference between parents on dominance attitude) constituted the second group of *Ss*. These two groups were likewise nearly identical in average age and average length of attendance in the nursery school. To test the related hypothesis, the mean time

(in minutes) spent in adult contact was determined for each group, and the difference between group means was tested for significance by the t test. Summary data for this analysis appear in Table 2.

The results support the second hypothesis ($t = 3.2$, $p < .05$, two-tailed test). Children with parents reflecting large differences on a dominance attitude scale displayed more adult avoidance behavior than children with parents reflecting small differences on a dominance attitude scale.

TABLE 2

SUMMARY OF CHILDREN'S AVOIDANCE BEHAVIOR: GROUPS DETERMINED BY DIFFERENCES BETWEEN PARENTAL DOMINANCE SCORES

Group	*Mean* (min.)	*SD*	*t*
I Children of parents with large differences ($N=4$)	5.8	1.1	3.2*
II Children of parents with small differences ($N=4$)	15.0	4.9	

* $p < .05$ (two-tailed test).

DISCUSSION

Two pertinent questions may be raised concerning the first hypothesis. The first question regards the sex of the parents. Are the dominance attitudes of mothers better predictors of their children's avoidance behavior than the dominance attitudes of fathers? The second question regards the sex pairing of parent and child. For example, are the dominance attitudes of mothers better predictors of avoidance behavior for daughters than for sons? These questions are discussed below.

An analysis of the data revealed no difference in predictive value between mothers' dominance attitudes and fathers' dominance attitudes. In both cases, the difference in avoidance scores between children with parents in the first and fourth quartiles and children with parents in the second and third quartiles was in the predicted direction and significant at the .10 level (two-tailed test). As indicated in the results section, the difference is significant at the .05 level when the dominance attitude scores are combined for parents. A further analysis also failed to disclose a closer parent attitude - child behavior relationship when parent and child are paired by sex. The dominance attitudes of both mother and father apparently affected the child about the same irrespective of the sex of the child.

A similar breakdown was made on the second hypothesis. Again the predictive value of dominance attitude scores was not enhanced either when the sex of the parents was considered separately or when pairings were made between parent and like-sex child.

The data were inspected to determine if children of parents reflecting highest difference scores on the attitude scale (Group I of Table 2) tended

also to be children of parents scoring at the extremes on the scale (Group I of Table 1). The inspection revealed that only one of the four children of parents in the high difference group also had parents in the extremes group. This relative lack of overlap suggests that our two hypotheses have been fairly independently confirmed by our data.

The present study supports the usefulness of the USC Parent Attitude Survey in differentiating between parents of normal children. In spite of the relatively small sample size, and the relatively homogeneous character of the sample, the scale was sufficiently sensitive to discriminate between groups at the .05 level of confidence on the two-tailed test.

Summary

Two hypotheses were tested regarding the relationship between parental attitudes and children's behavior. The first hypothesis stated that nursery school children of parents scoring either high or low on dominance attitude will avoid adult contacts in a free-activity setting to a greater degree than nursery school children of parents scoring in the intermediate range on dominance attitude. The second hypothesis stated that nursery school children of parents reflecting large differences in dominance attitudes will avoid adult contacts in a free-activity setting to a greater degree than nursery school children of parents reflecting small differences in dominance attitudes.

The sample consisted of 16 children, of equal sex proportion, with an average age of four years, three months, and their parents. The USC Parent Attitude Survey was administered to the parents, and the Dominance subscale served as the predictor variable. Five-minute time samples, totalling 50 minutes of observation in all, were made on each child. The time spent by a child in adult contact during the observation period served as the dependent variable.

The results supported the curvilinear relationship predicted in the first hypothesis at the .05 level (two-tailed test). The results also supported the second hypothesis at the .05 level. Further breakdowns of the data in terms of the sex of the parent or pairing the child to the like-sex parent did not increase the predictive power of the dominance scale. The combined attitudes for both parents were more discriminating than the attitudes for either parent considered separately.

References

1. Anderson, H., Hanley, C., & Hurley, J. Personality development in infancy and the preschool years. *Rev. educ. Res.*, 1955, 25, 453-468.
2. Ausubel, D. P., Balthazar, E. E., *et al.* Perceived parent attitudes as determiners of children's ego structure. *Child Develpm.*, 1954, 25, 173-184.

3. Ayer, M. E., & Bernreuter, R. G. A study of the relationship between discipline and personality traits in little children. *J. genet. Psychol.*, 1937, 50, 165-170.

4. Baldwin, A. L. Child psychology. In P. R. Farnsworth (Ed.), *Annual review of psychology*. Stanford, Calif.: Annual Reviews, 1956. Pp. 259-283.

5. Bornston, F., & Coleman, J. The relationship between certain parents' attitudes toward child rearing and the direction of aggression of their young adult offspring. *J. clin Psychol.*, 1956, 12, 41-44.

6. Champney, H. The variables of parent behavior. *J. abnorm. soc. Psychol.*, 1941, 36, 525-542.

7. Gordon, J. The validity of Shoben's Parent Attitude Survey. *J. clin. Psychol.*, 1957, 13, 154-156.

8. Koch, H. L., The relation of certain family constellation characteristics and attitudes of children toward adults. *Child Develpm.*, 1955, 26, 13-40.

9. Lorr, M., & Jenkins, R. Three factors in parent behavior. *J. consult. Psychol.*, 1953, 17, 306-308.

10. Mummery, D. V. Family backgrounds of assertive and non-assertive children. *Child Develpm.*, 1954, 25, 63-80.

11. Radke, M. J. The relation of parental authority to children's behavior and attitudes. *Univer. Minn. Child Welf. Monogr.*, 1946, No. 22.

12. Read, K. H. Parents' expressed attitudes and children's behavior. *J. consult. Psychol.*, 1945, 9, 95-100.

13. Sears, R., *et al.* Some child-rearing antecedents of aggression and dependency in young children. *Genet. Psychol. Monogr.*, 1953, 47, 135-236.

14. Sewell, W. H., Mussen, P. H., & Harris, C. W. Relationship among child training practices. *Amer. sociol. Rev.*, 1955, 20, 137-148.

15. Shoben, E. J. The assessment of parental attitudes in relation to child adjustment. *Genet. Psychol. Monogr.*, 1949, 39, 101-148.

A STUDY OF THE VALIDITY OF PARENT ATTITUDE MEASUREMENT

DONALD A. LETON
University of California at Los Angeles

The influence of parents' attitudes on the social and emotional development of children has long been recognized. Extensive programs of parent education are being conducted through the schools, clinics, and hospitals; yet seldom, if ever, are these programs evaluated in terms of the changes produced in parents' attitudes or in children's adjustment. The present-day confusion among parents on such important variables as permissiveness, control, and discipline seems to reflect the lack of definitive research in this area.

The most extensive work in this area has been in the development of the Parent Behavior Rating Scale at the Fels Research Institute by Champney (3). This was subjected to further analysis by Baldwin, Kalhorn, and Breese (1) who have published a manual for its use. Shoben (6) and Wheeler (7) have constructed inventories for the measurement of parent attitudes and have carried out subsequent standardization and validation studies. Radke's (5) study on parental authority and Bornston's research (2) on children's aggression represent two attempts to study certain aspects of parent-child relationships. A recent study by Gordon (4) tested the validity of Shoben's scale with mothers of deaf children. Although this is not a comprehensive review of the literature on parent attitudes, it is representative of the research which has been done.

PROBLEM

Most personality and attitude scales are validated through the use of clinical or criterion groups. The items subsequently included in the scale are chosen for their power to discriminate between the criterion groups. This method does not establish item or test validity for "normal" test populations, i.e., for those subjects who would not ordinarily be classified into the criterion groups.

CHILD DEVELOPMENT, Vol. 29, No. 4 (December, 1958)

The present study represents an investigation into the validity of parent attitude measurement. The two instruments which were used in the experiment were Shoben's Parent Attitude Survey (PAS) and the writer's revision of the Minnesota Teacher Attitude Inventory (MTAI) for parents.[1] Both of these inventories were originally validated through the use of criterion groups. The attitudes of parents of problem children and of parents of nonproblem children were contrasted in the validation of the PAS. The validation studies of the MTAI contrasted teachers with desirable and undesirable teacher-pupil relationships.

The revision of the MTAI for the measurement of parent attitudes (PAI) in this study was accomplished as follows: Of the 150 items in the scale, it was found that 118 could readily be indentified as "attitudes toward children." These items were reworded by changing "pupils" to "children," "teacher" to "parent," etc. The items were then keyed and scored in the same manner as for teachers.

Procedure

The PAS and the PAI were administered to a representative sample of parents in a suburb of Los Angeles, California,[2] obtained as follows: Alphabetical listings of the family names were obtained for all of the classrooms from kindergarten through the eighth grade at the two cooperating schools. Every fifth name was then selected from these lists to provide a sample of five families from each classroom, or approximately 20 at each grade level. The selection of names from the alphabetical listings varied with each grade level, e.g., for the first grade, 1st, 6th, 11th . . ., and for the second grade, 2nd, 7th, 12th, . . . name from each list. This procedure did not provide a sample of the total population of parents since no attempt was made to include parents whose children had not yet attained school age, nor those whose children had already completed junior high school.

In each family both parents were requested to respond to the inventories. The inventory booklets, answer sheets, and questionnaires were distributed to the homes by room mothers of the P.T.A. Parents were requested to respond to the inventories independently and asked not to discuss or compare responses until they had returned the materials. There were 283 parents who participated in the survey, of which 248, or 88 per cent, were complete pairs (*see* Table 1). The percentage of returns was approximately 78 per cent.

The mean age of the mothers in the sample was 39.2 (SD, 9.2), that of the fathers, 39.1 (SD, 6.9). The average educational attainment of these

[1] Appreciation is expressed for permission from the authors to use the Parent Attitude Survey and the Minnesota Teacher Attitude Inventory in this study.

[2] The writer is indebted to the Parent-Teacher Associations of Richmond Elementary School and the Junior High School of El Segundo, California, for their approval and assistance in conducting this study.

TABLE 1

CLASSIFICATION OF RESPONDENTS

Category	*Number*	*Per Cent*
Mothers responded, fathers nonrespondent	17	6
Mothers responded, no father present in home	6	2
Fathers responded, mothers nonrespondent or test data invalid	4	1
Mother-father pairs, test data complete (124 pairs)	248	88
Mother-father pairs, test data incomplete	8	3
Total	283	100

parents exceeded the thirteenth year (13.2) and ranged from the seventh grade level to the Ph.D. The average number of children was 2.9, with a range from 1 to 5.

RESULTS

It was first determined whether a significant relationship existed between the distributions of attitude scores for the two instruments. The correlation coefficient between the PAS and the PAI scores was found to be —.06. The negative coefficient is appropriate since poor attitudes are indicated by high scores on the PAS and by low scores on the PAI, and favorable attitudes, conversely. The magnitude of the relationship, however, is not significant and it is concluded that the two instruments are essentially unrelated in their measurement of parent attitudes.

The correlation between mother and father scores for the 124 pairs on the PAS was .43, on the PAI, .42, both coefficients significant at the .01 level.

It was next determined whether mothers had more favorable attitudes toward children than did fathers. The summary statistics are shown in Table 2. The differences between mothers' and fathers' scores on the two

TABLE 2

MEANS AND VARIANCES FOR PAS AND PAI DISTRIBUTIONS
(124 Mother-Father Pairs)

Instrument	MEAN		*Mean Difference*	VARIANCE	
	Mothers	*Fathers*		*Mothers*	*Fathers*
PAS	329.7	332.5	—2.8	491	366
PAI	78.5*	73.9*	—4.6	585	685

* In order to eliminate negative scores, a constant of 75 was added to each PAI score.

inventories were tested by a simple analysis of variance. The assumption of homogeneity of variance was found to be tenable. Although the mean differences were in favor of mothers on both instruments, the magnitude of the differences was not significant.

A total of 404 children were represented in the study, of whom 311 were of school age. It was possible to obtain adjustment ratings for 239 of the school-age children from their teachers. Ratings were not obtained for children in high school, or in private or parochial schools. The distribution of ratings is given in Table 3.

TABLE 3

DISTRIBUTION OF ADJUSTMENT RATINGS

Rating	*Personal, Social, and Emotional Adjustment*	*N*
1	Excellent	36
2	Very good	71
3	Average to above average	61
4	Average to below average	36
5	Barely adequate	18
6	Poor or unsatisfactory	17
	Total	239

The mean adjustment rating for the 239 children was 2.91. There was no check on the accuracy or reliability of the teachers' ratings, and for this reason only the extreme ratings were used in further analyses. For the 17 children who received "6" ratings there were 14 families represented; for the 36 children who received "1" ratings there were 25 families. There were no families in which two siblings were at opposite extremes of adjustment; the intermediate ratings for siblings were not considered. The comparison of the two parents groups is presented in Table 4. The differences

TABLE 4

ATTITUDES OF PARENTS OF CHILDREN WITH EXTREME ADJUSTMENT

Child Adjustment Rating	*N*	*Parent*	*N*	PAS		PAI	
				Mean	*SD*	*Mean*	*SD*
1	36	Mothers ..	25	328.3	41.0	91.3	24.8
6	17	Mothers ..	14	324.9	19.9	76.9	20.7
1	36	Fathers ..	25	327.2	21.8	79.8	23.2
6	17	Fathers ..	14	326.2	26.4	76.4	28.2

between the mean scores of parents of children with good adjustment ratings and of parents of children with poor adjustment ratings were not significant for either of the attitude inventories, as tested by analysis of variance.

The data permitted a test of a widely accepted concept in child development that parental agreement has a favorable effect on a child's adjustment and that parental disagreement has an unfavorable effect. The differences between the individual mother and father scores on the PAS and PAI were computed for the 39 pairs at the two extremes of child adjustment. The differences on both instruments were then combined into a single discrepancy score to determine whether the parents of the poorly adjusted children had a wider discrepancy, or "difference-in-attitude," score than did the parents whose children received the best adjustment ratings.

It can be seen in Table 5 that there is a greater discrepancy between the attitude scores of parents of poorly adjusted children than between parents of well adjusted children. The difference is significant ($t = 6.23$, $p < .001$).

TABLE 5

ATTITUDE DISCREPANCY SCORES OF MOTHER-FATHER PAIRS

Child Adjustment Rating	*Number of Pairs*	Discrepancy Scores	
		Mean	*SD*
1	25	36.5	23.8
6	14	55.4	23.9

SUMMARY

Shoben's Parent Attitude Survey and a revision of the Minnesota Teacher Attitude Inventory were administered to a representative sample of parents of children in kindergarten through the eighth grade. A summary of the findings follows:

1. There was no significant relationship between scores on the two attitude inventories.

2. There was a significant similarity between mother and father attitudes toward children within a given family.

3. Although mothers obtained more favorable scores than did fathers on both instruments, the differences were not significant.

4. There was no significant difference between the attitude scores of parents whose children received ratings of "excellent adjustment" and those whose children received "poor adjustment" ratings.

5. There was a wider disagreement in attitudes between mothers and fathers of poorly adjusted children than between mothers and fathers of well adjusted children.

References

1. Baldwin, A. L., Kalhorn, J., & Breese, F. H. Patterns of parent behavior. *Psychol. Monogr.*, 1949, 63, No. 4.
2. Bornston, F. L., & Coleman, J. C. Relationship between certain parents' attitudes toward child rearing and the direction of aggression of their young adult offspring. *J. clin. Psychol.*, 1956, 12, 41-44.
3. Champney, H. The measurement of parent behavior. *Child Develpm.*, 1941, 12, 131-166.
4. Gordon, J. E. The validity of Shoben's Parent Attitude Survey. *J. clin. Psychol.*, 1957, 13, 154-156.
5. Radke, Marian J. The relation of parental authority to children's behavior. *Univer. Minn. Inst. Child Welf. Monogr.*, 1946, No. 22.
6. Shoben, E. J. The assessment of parental attitudes in relation to child adjustment. *Genet. Psychol. Monogr.*, 1949, 39, 101-148.
7. Wheeler, M. E. Preliminary standardization of a parent-child relationship scale. *Dissertation Abstracts*, 1955, 15.

DEVELOPMENTAL STATUS OF TWO GROUPS OF INFANTS RELEASED FOR ADOPTION[1]

D. Bruce Gardner *and* Marybelle K. Swiger
Child Development Research Laboratories, Iowa State College

A research project undertaken by the Child Development Research Laboratories in 1955 had as its central objective the investigation of the effects of home management house experience on the development of infants. A discussion of the theoretical bases of the research has been presented earlier (3), and will not be included here.

In the over-all research project, various measurements of development and adjustment are being made at each of seven specified age levels during the first two years of life. Infants who live in a college home management house are compared, with respect to these measurements, with infants who live in private foster homes, and also with infants who live in their own parental homes from the time of birth. In order to make valid and meaningful comparisons of these three groups of infants, and relate these comparisons to the experimental conditions which define each of the groups, it is of fundamental importance to establish the degree of comparability of the groups before they are exposed to the unusual mothering experiences. This report describes the developmental status of the subjects at the time they were first inducted into the project, prior to being subjected to the experimental conditions under investigation.

There are obvious practical justifications for making the comparisons reported here, for the reasons explained above. In addition, there are possibly some theoretical implications which go beyond the demands of our project. The fundamental question of differences in genetic endowment

[1] Journal Paper No. J-3447 of the Iowa Agricultural and Home Economics Experiment Station, Ames, Iowa, Project No. 1299. Acknowledgement is due the Elizabeth McCormick Memorial Fund for partial financial support, and the staff of the Iowa Children's Home Society for their cooperation.

between infants whose mothers release them for adoption and infants who are born into complete families is raised here. The existence of constitutional differences as a function of legitimacy of birth, and the degree to which any such possible differences may be modifiable by early mothering experiences, will remain open questions in the absence of definitive research on the specific variable of legitimacy. However, the manner of selecting subjects for the groups in this study, and the over-all experimental design involved, may enable us to make some tentative statements on the role of genetic factors in the development of adopted children.

Description of the Subjects

In this report Group A includes 37 infants who subsequently are subjected to residence in a home management house for a period of one college quarter, or about three months, prior to adoption. Group B includes 44 infants of comparable background to those in Group A, but who reside in private homes with foster parents for a period of time before being adopted. Infants in both of these groups are born to unmarried mothers and released for legal adoption purposes. The selection of infants for the home management house was made by the staff of the Iowa Children's Home Society, which agency supervised the placement and adoption of the infants. The only basis of discrimination between Groups A and B was the practical consideration of availability of infants at the time each college quarter began.

A third group, C, is made up of 47 infants who were born into complete families and who will, presumably, be subjected only to "normal" mothering experiences in their own natural homes.

All of the subjects were born in the State of Iowa, and with few exceptions they were born in Des Moines hospitals. Infants with any obvious congenital defects such as spastic paralysis, malformations of limbs or body organs, micro- or hydrocephaly, or serious sensory defects, were excluded. Only full-term infants were accepted, the criterion for "full-term" being a minimum birth weight of 5½ pounds.

Table 1

CHRONOLOGICAL AGE IN DAYS AT TIME OF EXAMINATION

	Males			Females			Total		
	N	*Range*	*Mean*	*N*	*Range*	*Mean*	*N*	*Range*	*Mean*
Group A	21	7–59	27.1	16	4–72	29.0	37	4–72	27.9
Group B	21	8–64	30.6	23	7–75	27.3	44	7–75	28.9
Group C	25	21–92	39.5	22	18–57	34.2	47	18–92	37.0
Total	67	7–92	32.9	61	4–75	30.2	128	4–92	31.6

The sex distribution within groups was controlled as rigidly as practicable within the limits of availability of infants, and the desirability of including as many subjects as possible. However, the rigid control of this variable was thought to be a less vital consideration than maintenance of reasonably good sized groups, since sex differences were not the prime variable under consideration. Table 1 reveals the degree to which the three groups and the total were comparable with respect to relative proportions of males and females.

In attempting to determine the comparability of the three groups, a most important consideration is that of the age at which the measurements were taken. Efforts were made to control the examination age, by limiting the enrollment of subjects to those under three months of age, and by attempting to test them as near as possible to the 30th day following birth. For reasons of scheduling and practical considerations in the placement of infants in the home management houses and foster homes, it was impossible to control this variable as well as would have been desired. Table 1 provides data showing the mean ages and age ranges at time of examination, for the three groups and also for sexes.

The mean ages for males and females were not significantly different, either within any group or for the combined groups. However, a comparison of the three groups without regard to sex revealed that Group C infants were significantly older than were those in the other two groups ($F = 3.686$, significant at the .05 level with 2 and 127 *df*). This difference necessitated a statistical control of the age variable in order to analyze the data for comparative developmental status of the three groups.

Table 2 provides information on two measurements of gross body development at birth, those of weight and length. The sex difference in birth length approaches significance at the .05 level ($F = 3.788$ with 1 and 123 *df*). Analysis of variance was used to test the significance of differences among groups, as well as between sexes, for both weight and length at birth. However, none of the *F* ratios, other than that for sex difference in length, was near significance at the .05 level, and they are not shown.

TABLE 2

PHYSICAL MEASUREMENTS AT BIRTH

	MALES			FEMALES			TOTAL		
	N	*Weight* (kg.)	*Length* (cm.)	*N*	*Weight* (kg.)	*Length* (cm.)	*N*	*Weight* (kg.)	*Length* (cm.)
Group A	21	3.40	50.84	15	3.31	50.67	36	3.36	50.77
Group B	20	3.41	51.98	22	3.44	51.35	42	3.42	51.65
Group C	24	3.51	51.57	22	3.20	49.68	46	3.37	50.66
Total	65	3.44	51.46	59	3.32	50.55	124	3.39	51.03

Similarly, interaction effects of sexes and groups were examined with the *F* test, and were found to be negligible for both variables.

A thorough medical examination by the consulting pediatrician was made prior to induction of each infant into the research. Since this screening examination eliminated any infants with marked pathological conditions, the net effect was to restrict the range of variation in physical status among the subjects. The nature of the medical examination reports, plus the fact that the average age of infants in Groups A and B (10 days) was considerably lower than that for Group C (31 days) when the medical records were taken, made statistical comparison of the groups difficult and hazardous.

A simple tabulation of positive medical findings in each of the areas included in the pediatric report (including body temperature, skin, sensory organs, nose and throat, abdomen, heart, lungs, genitals, body posture, and general nutritional status) revealed the following incidence of anomalies: Group A, 46; Group B, 40; Group C, 33. In terms of percentages, Group A, which included 30 per cent of the cases, accounted for 39 per cent of the anomalies. Group B, with 35 per cent of the cases, contributed 33 per cent of the anomalies. Group C, with 35 per cent of the cases, contributed 28 per cent of the anomalies.

Unfortunately, statistical tests of differences for these proportions would be relatively meaningless, for at least two reasons. First, the wide range of items listed above and lumped together under the heading of "anomalies" cut across many dimensions of observation and measurement, and they are certainly not of equal significance medically. Second, such variables as general nutritional status, which was more frequently noted as a positive finding in Groups A and B than in Group C, are probably a function of age. In general, there is nothing in the medical data to indicate that one of the groups is significantly inferior or superior to the others.

It might be expected that differences in developmental status among the three groups could arise from hereditary variation, if it could be assumed that there are some heritable characteristics associated with out-of-wedlock birth. However, there seems to be little justification for making any such assumption in the absence of any evidence that the tendency to out-of-wedlock parentage is associated with genetic structure. Lacking such evidence, we are assuming that genetic effects on developmental status are distributed no differently in one group than in the others. The data reported here are, in one sense, a test of the validity of this assumption, with respect to a limited number of measurements of developmental status taken within a relatively narrow range of chronological age.[2]

---

[2] Admittedly this assumption may not be fully warranted since information on parentage in Groups A and B is meager. Furthermore, the age of mothers at time of subjects' birth is lower, on the average, for Groups A and B than for Group C. The extent to which this factor may have influenced the constitutional make-up of the subjects cannot easily be determined.

Measuring Instruments and Examination Procedure

The variables of physical and psychological status are represented by a variety of specific measurements. Anthropometric measurements include body weight, length, head circumference, chest circumference, calf circumference, bi-iliac (hip) width, and subcutaneous leg tissue. Standard measuring instruments were used to obtain these data. Steel tapes, a standard measuring board, and basket-type scales were used to obtain circumferences, length, and weight respectively. Bi-iliac width was measured with large sliding calipers; small spreading calipers were used to measure subcutaneous tissue in the left leg. Calf circumference was also measured on the left leg.

Body measurements were taken by a nurse and an assistant, both of whom received preparatory instruction in techniques of anthropometric measurements with infants. Each measurement was made twice, the second time by the assistant, and the average of the two readings was computed. If at any time a sizable discrepancy was noted between the two readings, a third one was taken and averaged with the nearer of the first two.

The Gesell Developmental Schedules were administered by a psychologist, following the procedure outlined by Gesell and Amatruda (2), and using the standard examination materials. The scales were designed to estimate developmental status in each of four areas (language, motor, adaptive, and personal-social), as well as providing a total score. Since the number of items in each area is small, and the precise meaning of the areas somewhat open to question, only a general developmental level was computed for purposes of this report.[3]

The California Infant Scale of Motor Development was also administered by a psychologist. This scale was developed by Bayley (1) for measuring motor skills and coordination. Each item is age-scaled, and the items range from .2 months to 54.5 months. A point score is computed, based on the actual number of items passed. This is then converted to a sigma score, on the basis of mean and standard deviation figures provided for each age level by Bayley.

Each subject was examined individually, and 30 to 40 minutes were required to complete the psychological and anthropometric measurements. Subjective notes of the examination behavior of the infant were recorded by each of the two psychologists present at the examination. The sequence

[3] A somewhat more precise quantification of scores on the Developmental Schedules than that suggested by Gesell and Amatruda (2) was desired for purposes of statistical analysis. Therefore, the scales were treated as if they were a psychometric test, and the child's performance level was computed as follows: key age was determined by noting the highest test level at which over half the items were passed. This estimate of developmental age was corrected by subtracting a number of days proportionate to the number of failures at that level and at lower levels. If items were passed at higher levels, a number of days of developmental age proportionate to the number of passes was added to obtain developmental age level. This figure was divided by chronological age and multiplied by 100 to obtain the general developmental quotient.

of tests was always the same: the Developmental Schedules were followed by the Scale of Motor Development, after which the anthropometric measurements were taken. An examination room at the Iowa Children's Home Society was used for Groups A and B, and an examining room at the pediatric clinic was used, by the same team of examiners, for Group C. Each infant was accompanied in the testing room by his mother, or foster mother. The length of time required to complete the induction of all subjects into the project was approximately one year.

Results

The basic findings, in the form of mean scores in each of the variables, are presented in Table 3. Tests of significance of differences between sexes revealed that the mean scores favored males in body length ($F = 5.205$, significant at the .05 level), and in head circumference ($F = 5.844$, significant at the .05 level). None of the other six differences was significant at that level; however, the difference in weight, which also favored males, was significant at the .10 level ($F = 3.053$). The remaining anthropometric measurements, with the exception of leg tissue, favored males but the differences were small and not statistically reliable.

Table 3

MEAN SCORES AND MEASUREMENTS FOR GROUPS AND SEXES

Variable	Group A		Group B		Group C		Males		Females		Total	
	Mean	*N*	*Mean*	*N*	*Mean*	*N*	*Mean*	*N*	*Mean*	*N*	*Mean*	*N*
Weight (kg.)	3.94	37	4.01	44	4.48	46	4.29	66	4.02	61	4.16	127
Length (cm.)	53.44	37	54.09	44	54.77	46	54.79	66	53.45	61	54.15	127
Head (cm.)	36.66	35	36.46	43	36.94	44	37.07	62	36.30	60	36.69	122
Chest (cm.)	34.09	36	34.26	43	35.19	45	34.86	64	34.21	60	34.55	124
Calf (cm.)	11.83	35	12.04	43	12.78	44	12.27	62	12.22	60	12.25	122
Hip (cm.)	8.59	37	8.68	43	8.84	46	8.76	66	8.67	60	8.72	126
Leg tis. (mm.)	15.25	37	15.71	39	17.38	44	16.03	62	16.34	58	16.18	120
Motor (sigma)	1.48	35	1.41	44	0.58	46	1.06	66	1.20	59	1.13	125
DQ	135	37	136	44	93	47	114	67	126	61	120	128

A statistical comparison of the three groups with respect to physical measurements necessarily had to take into account the discrepancy in chronological age, since it might be assumed that these variables are correlated with age. Therefore, analysis of covariance was used to test for significance of differences among group means adjusted for age. Table 4 presents the basic data derived through the covariance analyses.

The consistently significant correlation of each of the variables with age is one indication of the justification for resorting to the covariance analyses. Comparisons of the adjusted mean scores for groups, shown in Table 4, with the unadjusted mean scores reported in Table 3 reveals the tendency for the age adjustment to bring the mean scores for groups much closer to each other. The *F* test here represents the ratio of the adjusted mean square for variance between groups to the adjusted mean square for variance within groups. Only in the case of calf circumference is the *F* ratio significant ($F = 3.113$, significant at the .05 level with 2 and 120 *df*).

TABLE 4

MEAN SCORES AND MEASUREMENTS ADJUSTED FOR AGE

Variable	*r**	*Group A*	*Group B*	*Group C*	*F*
Weight (kg.)	.68	4.06	4.10	4.29	1.494
Length (cm.)	.66	53.90	54.43	54.14	.430
Head (cm.)	.69	36.86	36.66	36.67	.288
Chest (cm.)	.77	34.39	34.47	34.69	.483
Calf (cm.)	.66	11.96	12.14	12.53	3.113
Hip (cm.)	.65	8.72	8.74	8.67	.098
Leg tis. (mm.)	.57	15.59	16.18	16.76	1.530
Motor (sigma)	—.27	1.17	1.32	0.64	1.154
DQ	—.64	126	129	107	2.070

* Correlation of each variable with chronological age.

Statistical summaries of the scores obtained with the California Infant Scale of Motor Development and the Gesell Developmental Schedules are also presented in Table 4. It should be noted that in both of these sets of data scores are reported in the form of derived indices, sigma score and developmental quotient respectively. Since the derivation of these scores from raw scores involves computing a ratio of performance to chronological age, the use of the derived index should have the effect of controlling the age variable. However, several characteristics of the data suggest the presence of a negative correlation between derived scores and chronological age. After computing the correlation coefficients and noting the extent of this relationship, it seemed appropriate again to make use of the covariance technique to control the age variable.

Adjusted mean scores for motor development and developmental quotient are shown in Table 4. The *F* ratios indicate that the differences are not significant for either variable.

Discussion

The one measurement in which significant differences among groups remain after the age adjustment is made, that of calf circumference, is not easily interpreted in relation to the rest of the data. The absolute difference between any two means is not great: the largest, between Groups A and C, is slightly more than half of one centimeter. This could not be said to be of fundamental significance as a basis for distinguishing between two groups. Presumably, in any large series of statistical tests of differences among given groups, the chances are fairly good that a "significant" F ratio may occur in the absence of any very important real difference.

Nevertheless, the primary purpose of the analyses reported here was to establish a base line of comparison of the three groups prior to experimental treatment in the larger longitudinal study. Since this difference in calf circumference, however small, does appear to be associated with the selection of subjects for the groups, it should be of interest to note, in subsequent analyses of data obtained at older ages, whether this difference remains over a period of time allowing for maturation and more extensive use of the body in walking and other motor activity.

So far as the remaining variables are concerned, it may safely be assumed that the three groups are not fundamentally different with respect to the variables we have measured. The logic implicit in the statistical tests employed here does not allow us to say that the subjects in the different groups are all members, drawn at random, of some larger population. Speaking conservatively, however, it is evident that no one group differs consistently or systematically from the other two.

There are both practical and theoretical implications of these findings. From the standpoint of the larger research project, this serves to increase our confidence in the validity of comparisons to be made following experimental treatment of Groups A and B. The question may be raised, of course, about underlying genetic differences which might not have had time to appear through normal maturation processes during the first few weeks after birth. The best answer to this question, so far as the longitudinal design is concerned, will probably become available through a careful use of the foster home group data. Since these subjects are selected in the same manner and from the same sources as are the home management house infants, any underlying differences in genetic endowment between the home management house group (A) and the control group (C) should also appear between the foster home group (B) and the control group (C).

One interesting aspect of the results certainly is the striking negative correlation between developmental quotient and chronological age. A possible explanation for this finding lies in the nature of the Gesell Schedules and the normative data available on these scales for the early weeks of infancy. For the four-week level, the earliest age at which norms are available,

the items do not appear to discriminate well among infants whose chronological ages range from 0 to 30 days. That is, the infant who is one week old frequently is observed to "pass" as many of the four-week items as the infant whose age is three or four weeks. This failure to discriminate finely results in a distortion which is mathematically amplified in the process of computing a DQ score with such young subjects. The tendency of the youngest subjects to receive disproportionately high DQ scores undoubtedly contributes to the strong negative relationship between CA and DQ. It would seem appropriate to suggest that research workers using the Gesell Schedules with very young infants might, if it is important to compute DQ scores, add a constant both to the numerator and the denominator in the DQ formula. The addition of, say, 60 days should tend to reduce this apparent distortion considerably.

Summary

The longitudinal investigation of the effects of noncontinuous mothering on human development necessitated comparison of the subjects prior to exposure to the different kinds of mothering under investigation. This report describes the developmental status of three groups of infants at the time they were first inducted into the research, prior to any experimental "treatment."

Two groups of infants, A and B, were born to unmarried mothers and released by them for adoption. A third group, C, was composed of infants born into complete families under normal circumstances. The proportion of subjects of each sex in each group was kept relatively even. Efforts were made to examine all infants at or near one month of age; however, Group C infants were slightly older, on the average, than those in the other two groups. This difference made it necessary to apply statistical control to the age variable in order to compare the groups.

For each group, data were obtained on a series of anthropometric measurements, motor development as rated with the California Infant Scale of Motor Development, and general developmental status as rated with the Gesell Developmental Schedules.

Significant sex differences, favoring male subjects, were noted in body length and head circumference. Analysis of the data by groups revealed that Groups A, B, and C were roughly comparable in developmental status at the time of induction into the research. In only one measurement did a significant difference appear, that of calf circumference. It would seem as plausible to interpret this as a chance variation, which could occur in such a series of tests of significance, as to assume that it is a reliable index of fundamental differences among the groups, since there is little else in the data to support this specific finding.

Significant negative correlations of developmental quotients with chronological age were noted, and were attributed largely to the inadequate dis-

criminatory powers of the scales during the early weeks, as well as the very small CA factor used in the denominator of the DQ formula. A mathematical correction for the distortion resulting from these two conditions was suggested.

References

1. Bayley, Nancy. The development of motor ability during the first three years. *Monogr. Soc. Res. Child Develpm.*, 1936, 1, No. 1.
2. Gesell, A., & Amatruda, Catherine. *Developmental diagnosis.* (2nd Ed.) New York: Hoeber, 1954.
3. Pease, Damaris, & Gardner, D. B. Research on the effects of non-continuous mothering. *Child Develpm.*, 1958, 29, 141-148.

VARIABLES IN EARLY DISCRIMINATION LEARNING: II. MODE OF RESPONSE AND STIMULUS DIFFERENCE IN THE DISCRIMINATION OF TONAL FREQUENCIES[1]

WENDELL E. JEFFREY
University of California, Los Angeles

When one uses adult *Ss* in a discrimination experiment, the indicator response rarely causes any problems. One can ask that any of a number of types of judgments be made, in a number of different ways, without substantially increasing the difficulty of the task. With young children, however, the response required can make a critical difference in the speed of learning, or indeed, may make the difference between learning and not learning. For example, Jeffrey (3) demonstrated that children who could not learn to apply distinctive labels to stimuli differing in spatial orientation could nevertheless learn to press buttons toward which the stimuli (stick figures) could be said to be pointing. It seemed important to investigate further the various factors that contribute to this early discrimination learning. The spatial discrimination problem previously used was a relatively easy one. In order to get more reliable measures of the contributing factors, it was advisable to find a stimulus dimension that was more difficult and one that was also more readily manipulable.

The tonal frequency dimension (pitch) met these criteria. Preliminary research indicated that kindergarten children had difficulty in responding differentially, and that the words, "high" and "low," meant nothing to children when applied to musical notes, indicating that the stimuli were

[1] This research was supported by grant M-1498 from the National Institute of Mental Health, Public Health Service, and by grant G-4414 from the National Science Foundation. I wish to express my appreciation to Mrs. Eleanore Cornberg of the University School Kindergarten for her cooperation in this research.

not already verbally predifferentiated. Although adult humans tend to associate low tones with left, whether the button on the left was to be associated with a high or low tone made no difference to the children. Nevertheless, it seemed reasonable to believe that a response might be found that would be readily learned and possibly facilitative in the acquisition of those responses previously demonstrated to be very difficult to learn.

A response to pitch that seemed to have some of the feedback characteristics of the button pressing response to the stick figures in the research previously noted was that of singing a matching tone. Therefore, this task was chosen for one group of *Ss*. Because it was found that it is difficult for children to sing a particular note at will, another group was trained on an analogous task, viz., to produce matching tones on the piano as their indicator response. Button pressing, with buttons oriented to the right and left of *S*, provided a control task performed by a third group and served as a criterion test task after training with singing or with the piano.

Inasmuch as the discrimination of pitch proved to be a very difficult task, it was thought that the variable of stimulus difference should also be manipulated. The data of psychophysics and studies of stimulus generalization indicate that when stimulus differences are large discrimination learning is likely to be easy. There is, however, little information regarding the transfer of training from learning an easy discrimination to learning a more difficult one, when difficulty is defined in terms of separation along a stimulus continuum.

It could be said that the first responses a child has to learn in life are of the sort that define or orient the child to the continuum (1) and that gross stimulus differences help achieve this end. Once the continuum is defined, subsequent learning with regard to stimuli within the continuum would be expected to be facilitated. The present experiment first presented the children of the experimental groups with an opportunity to define the pitch continuum, then measured their success by means of a task otherwise beyond their ability, and finally made the latter task even more difficult by requiring closer discrimination. Two hypotheses were being tested: (a) that practice on the two separate pretraining tasks involving the production of tones to match the stimulus tones would transfer to the difficult task of pressing the left button after a low stimulus tone and the right button after a high stimulus tone; and (b) that training with a pair of tones separated by 1024 cycles would facilitate subsequent discrimination of a pair of tones separated by only 128 cycles.

Method

Subjects

The *Ss* were 21 children, 11 boys and 10 girls, from the University of California Senior Kindergarten. They were divided into three groups of seven each. The mean ages were 5 years, 6 months for Groups I and III

and 5 years, 7 months for Group II. The range for Group I was 5-2 to 5-10, for Group II, 5-4 to 6-0, and for Group III, 5-4 to 5-11.

Apparatus

Two tape recorders were used, one to present children's music, which was an incentive and reinforcer, and one to present the discriminative stimuli. Further apparatus programmed push buttons in such a way that if the correct button was pressed within three seconds following the occurrence of a discriminative stimulus the music continued to play. If the incorrect button was pressed, the music terminated. The music could be heard again only upon making a correct response following the presentation of the next discriminative stimulus. The response buttons were approximately 14 inches apart on a small board that could be placed on a table in front of *S*, or on the floor to be operated with *E*'s feet. The discriminative stimuli and the music were delivered through separate earphones worn by *S*. The tone was sufficiently loud to cause no confusion when it occurred along with the music. The results of the piano response group, presented below, confirm this.

Stimuli

Two series of piano tones were recorded. The first series, Series A, consisted of two tones, C_1 and G^3 (128 and 1152 cps.), recorded serially at 10-second intervals in the order H L H H L H L L. A second series, Series B, consisting of the tones, C and G (256 and 384 cps.), was recorded in the same manner.

Procedure

There were three groups of *S*s. Group I was trained only on the button pressing response; Group II was required to match the discriminative stimuli vocally prior to being tested on the button pressing response; and Group III learned to produce matching tones on the piano before being tested on the button pressing response. All of the above training and testing was done with Series A. Any *S* who met the criterion on the button response with Series A was tested for transfer to Series B. A criterion of 10 successive correct responses was used throughout.

Each *S* in Group I was brought into the experimental room and seated at a low table upon which were push buttons. Earphones were placed on his head and he was permitted to listen to music for approximately one minute. The music was then terminated and *S* was told that he would hear tones in the earphones and *E* sang representative tones and indicated which button was to be pressed for each tone in order to continue the music. The music was then turned on again and Series A (128 and 1152 cps.) was presented. Transfer of the button pressing response to Series B (256 and 384 cps.) stimuli was tested if the criterion of 10 successive correct discriminations on Series A was met within 72 trials. An earlier study indi-

cated that if *S*s had not learned in 72 trials they were not likely to learn even if they could be kept in the situation for a longer period.

The *S*s of Group II were instructed similarly to Group I except that they were asked to reproduce the Series A stimuli by singing or humming the tone that they had heard rather than pressing a specific button. Examples were given by *E* and the children were helped to move their voices to the correct tone by *E*'s singing along with them. On meeting the criterion on the singing response, these *S*s were presented with the button pressing task with the Series A stimuli. If they met the criterion of 10 successive correct responses on the button task, transfer of this response to Series B stimuli was also tested.

Each *S* in Group III was seated at a piano in the experimental room rather than at the table. After listening to the music, he was shown that two keys of the piano were marked with red tape, and was told that his task was to match the tone he heard through the earphones by pressing the correct key. As with the other groups, this task was demonstrated by *E*'s singing representative tones. Subsequent to meeting the criterion on the Series A stimuli, performance on the button response was tested. Those *S*s who met the criterion on this task were then presented with the Series B stimuli.

Results

Of first consideration in the results is the fact that only one of seven *S*s in Group I learned to press the buttons differentially to the Series A stimuli within 72 trials. One *S* in this group refused to continue after 48 trials. This failure to learn or to continue meant a score of zero or, more exactly, an infinite score. Thus, the data can be best summarized in a pass-fail contingency table (Table 1), with passing defined as meeting a criterion of 10 successive correct responses within the 72 trials.

Considering first the button pressing criterion responses to the Series A (128 and 1152 cps.) stimuli, only one *S* met the criterion in Group I, five *S*s met the criterion in Group II, and seven *S*s met the criterion in Group

Table 1

NUMBER OF CHILDREN MEETING CRITERION OF 10 SUCCESSIVE CORRECT DISCRIMINATIONS IN THE ALLOTED NUMBER OF TRIALS

	N	*Series A* $C_1 - G^3$	*Series B* C – G
Group I (buttons)	7	1	0
Group II (singing)	7	5	1
Group III (piano)	7	7	1

III. Group II *Ss* had had from a minimum of 24 to a maximum of 72 prior trials attempting a differential singing response. The median was 24 responses, four *Ss* seemingly being able to sing different notes reliably to the stimuli although not necessarily reproducing them exactly as presented. The general inability of the children in Group II to come even close to reproducing the stimuli made it difficult for *E* to score their responses. One *S* never produced differential tones that were in any way acceptable. The other *Ss* were reinforced for anything that sounded like a differential response in the hope of shaping their subsequent responses. Whether two observers, however, could have agreed on the acceptability of each response is not known. The fact remains that of the six *Ss* who, in *E*'s opinion, were learning to make differential singing responses, five subsequently learned to make the button pressing response.

Group III had from 10 to 24 trials making the piano responses prior to the test situation. Four of these *Ss* met the criterion with the piano immediately after instructions. The other three *Ss* took 9, 12, and 15 trials to meet the criterion with the piano. Transfer to the button response was perfect in all seven cases. Thus, for Group III, the criterion for the button pressing to Series A stimuli was reached by four *Ss* without experiencing failure and with only 10 pretraining trials. In the other three cases it took 19, 22, and 25 trials, 10 of these trials in each case being criterion trials for the piano response.

The differences in perfomance on the transfer task (button pressing) can best be evaluated by the Fisher exact probability test which tests the hypothesis as to whether two groups differ in the proportion with which individuals fall into two classifications. The difference between Groups I and III is significant with $p < .01$. The difference between I and II has an exact probability of .0512. Group II does not differ significantly from III. Thus, the hypothesis that such pretraining would be facilitative is supported. These findings are almost identical to those obtained in a preliminary study reported elsewhere (2).

The hypothesis regarding transfer from training with the Series A stimuli, involving a difference of three octaves and a fifth between stimuli, to the closer discrimination of the Series B stimuli, a difference of only a fifth, was not supported (Table 1). In Group I only one *S* learned the Series A discrimination, and he did not learn the Series B discrimination. In Group II, although five *Ss* learned the Series A discrimination only one *S* demonstrated any transfer to the Series B stimuli. In Group III, even though all *Ss* learned the button pressing response to Series A, again only one *S* showed transfer to the Series B stimuli.

Discussion

Impossible as it may seem, 5½-year-old children who can sing their favorite song (e.g., Davy Crockett) with quite acceptable intonation, are

not able to make a differential button pressing response to pitch stimuli without certain pretraining. They also do very poorly at singing tones that match presented stimuli even when given all possible aid by *E*. The degree to which the *S*s had difficulty with the button pressing task needs to be emphasized. The instructions were very specific and an *S*, after instructions, might make four or five correct responses in succession only to apparently suddenly become confused and no longer able to remember which button was appropriate for each tone. His performance then dropped to chance level. In greatest contrast to this performance is the performance of the piano response group which, on the whole, had very little difficulty making what in practically all aspects was an identical response except for the feedback it produced. This feedback apparently accounted then for not only faster learning on the piano response task but also the high positive transfer to the button pressing task. The fact that the piano response group learned so readily also indicates that the incentive conditions were sufficient.

There may be some question about the stability of this training, however, over longer periods of time. Therefore, before one can go very far in speculating about the processes involved, it is important to know whether the button pressing response will persist over a period of days, rather than just over a number of trials. If it is only temporary, it is highly probable that its permanence could be increased by further training. It is of theoretical importance to know, if more training is necessary to assure permanence, whether this added training would be more beneficial if done on the piano or with buttons.

In similar perceptual situations with children, psychologists have often been unable to determine whether failure to make a discrimination is due only to *S*'s inability to make an overt response or to his inability to make a discrimination of any kind. In this instance the success of the piano group suggests that *S*s are able to make the discrimination and therefore their failure to sing must result from the lack of differentiation of the singing response. On the other hand, it is to be noted that the button pressing group could not respond differentially. The question arises as to whether this group could discriminate but could not make the response or whether the basic perceptual discrimination occurred as the result of making the piano and singing responses. Such an hypothesis is supported by the fact that the singing group subsequently performed the button pressing response even though their singing responses had been painfully inaccurate. These results apparently may be explained in several ways. First, one might assume that the feedback of the singing and piano responses is important in some way. Certainly, it is hard to think of a tone or a tone discrimination without making a covert response. Both Mandler (5) and Kanfer (4) have accentuated the importance of such responses in discrimination learning, and to require a singing response should help implement this process. One would not necessarily predict, however, that the piano response would produce the appropriate mediational responses. A second hypothesis seems

more plausible here. One might assume that the correctness of a response is accentuated by the immediate comparison between stimulus tone and the response-produced tone. This would provide a relatively easy same-difference discrimination and should help accentuate the continuum involved as well as provide a more direct form of reinforcement. A third explanation, appropriate at least for the piano response group, is the fact that the piano response shortens the interstimulus interval by effectively extending each stimulus in time. The relevance of this variable is, of course, subject to test and work in this direction is now in progress.

Another source of difficulty in the discrimination may have arisen from the fact that there were common overtones between the two notes within each series, inasmuch as the piano was used as a stimulus source. This variable will also be assessed in future research.

Summary

In order to study some factors basic to early discrimination learning in children, the tonal frequency dimension was chosen as the source of stimuli. Both the type of indicator response and the difference between stimuli were varied.

The *Ss* were 5½-year-old children. The criterion task was to learn to push two buttons, oriented to the left and right, differentially to two tones separated by three octaves and a fifth. Group I had training only on this task. Prior to testing on the criterion task, Group II *Ss* were required to match the same two tones by singing, and Group III produced matching tones on a piano. Although *Ss* in Group II had difficulty making a match, they did seem to make differential responses, and when transferred to the button pressing criterion task, these *Ss* performed significantly better than those in Group I. The Group III *Ss*, who were required to produce matching tones on a piano, learned this quite readily and showed almost perfect transfer to the criterion task. Thus, Groups II and III differed significantly from Group I but not from each other.

All *Ss* who met the criterion on the button pressing task were subsequently required to make this same response to stimuli differing by only five-eighths of an octave. There was no evidence of transfer to these new stimuli in any of the groups. Suggestions for further research are discussed.

References

1. Ehrenfreund, D. An experimental test of the continuity theory of discrimination with pattern vision. *J. comp. Psychol.*, 1948, 41, 408-422.
2. Jeffrey, W. E. An experimental study of some factors in early discrimination learning. *Amer. Psychologist*, 1957, 12, 392. (Abstract)

3. JEFFREY, W. E. Variables in early discrimination learning: I. Motor responses in the training of a left-right discrimination. *Child Develpm.*, 1958, 29, 269-275.

4. KANFER, F. H. Perception: Identification and instrumental activity. *Psychol. Rev.*, 1956, 63, 317-329.

5. MANDLER, G. Response factors in human learning. *Psychol. Rev.*, 1954, 61, 235-244.

Child Development, one of three publications issued by the Society for Research in Child Development, Inc., is issued quarterly in March, June, September and December. The subscription price per year is $8.00 domestic, $8.50 foreign. Single issues are $2.50.

Child Development Abstracts and Bibliography is issued three times a year, two numbers in each issue. The subscription price per year is $6.00 domestic, $6.50 foreign. Single issues are $2.50.

Monographs of the Society for Research in Child Development, issued irregularly during the year, consist primarily of detailed critical reviews and seriatim studies. The subscription price per year is $6.00 domestic, $6.50 foreign. The price of individual monographs is variable.

The three publications of the Society are available at the special price of $16.00 per year, domestic; $17.50 per year, foreign.

A limited number of back issues of all publications is available. A mimeographed list may be had upon request.

Subscriptions, address changes, and business communications should be sent to CHILD DEVELOPMENT PUBLICATIONS, Purdue University, Lafayette, Indiana.

Please give four weeks' notice when changing your address, giving old and new addresses. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee forwarding postage. Other claims for undelivered copies must be made within four months of publication.

Printed and mailed by
THE ANTIOCH PRESS
Yellow Springs, Ohio

CURRENT MONOGRAPHS
of the Society for Research in Child Development, Inc.

Lester E. Harrell, Jr.
A COMPARISON OF THE DEVELOPMENT OF ORAL AND WRITTEN LANGUAGE IN SCHOOL-AGE CHILDREN 1957, $2.50

James J. Gallagher
A COMPARISON OF BRAIN-INJURED AND NON-BRAIN-INJURED MENTALLY RETARDED CHILDREN ON SEVERAL PSYCHOLOGICAL VARIABLES 1957, $2.50

Martin Lakin
PERSONALITY FACTORS IN MOTHERS OF EXCESSIVELY CRYING (COLICKY) INFANTS 1957, $1.75

Harriet Lange Rheingold
THE MODIFICATION OF SOCIAL RESPONSIVENESS IN INSTITUTIONAL BABIES 1956, $1.75

Richard M. Brandt and Hugh V. Perkins
RESEARCH EVALUATING A CHILD STUDY PROGRAM 1956, $2.75

Hilda Knobloch and Benjamin Pasamanick
A DEVELOPMENTAL QUESTIONNAIRE FOR INFANTS FORTY WEEKS OF AGE: AN EVALUATION 1956, $2.50

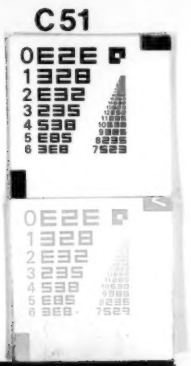
Wilton Marion Krogman
THE PHYSICAL GROWTH OF CHILDREN: AN APPRAISAL OF STUDIES 1950-1955 1956, $2.50

Jacob L. Gewirtz
THREE DETERMINANTS OF ATTENTION-SEEKING IN YOUNG CHILDREN 1956, $1.50

Lloyd M. Dunn and Rudolph J. Capobianco
STUDIES OF READING AND ARITHMETIC IN MENTALLY RETARDED BOYS 1956, $3.00

(A mimeographed list of all monographs available in the series since 1935 will be sent upon request.)

order from
CHILD DEVELOPMENT PUBLICATIONS
Purdue University, Lafayette, Indiana


C51